University of Southampton

Faculty of Arts and Humanities

Winchester School of Arts

Download First, Privacy Later: Exploring the exchange of smartphone usage data for access to app-based services in India.

DOI [10.5258/SOTON/D2482] https://doi.org/10.5258/SOTON/D2482

by

Ravi Dixit

ORCID ID 0000-0002-2794-8702

Thesis for the degree of Doctor of Philosophy

December 2022

University of Southampton

Abstract

Faculty of Arts and Humanities

Winchester School of Arts

Doctor of Philosophy

Download First, Privacy Later: Exploring the exchange of smartphone usage data for access to app-based services in India.

by

Ravi Dixit

This thesis reviewed scholarship in the field of media and cultural studies with specific focus on audience studies. Literature review was focused on key themes such as data practices and experiences, data privacy, advertising surveillance and issues around informed consent by users of digital platforms. The review identified a gap in research about user understanding of data exchange with digital platforms and experiences with personalised advertising in India. A mix of qualitative research methods were used to evaluate contemporary experiences of participants engaging with their personal data. This research explored the process of converting digital data trail into customised advertising for smartphone owners and their incentives to participate in this exchange of data for access to app-based services**.** Industry Experts working in media agencies and marketing function of advertisers were interviewed to understand various sources of digital data that are used for creating affinity audiences. Owners of smartphones in Mumbai and Delhi reflected on their awareness of automated data collection, trust threshold for sharing sensitive data and understanding of advertising technology. Participants discussed the affirmative nature of digital platforms and described services that were critical to their day-to-day functioning. Participants had a gradation of trust and all the apps needed to clear thresholds to be considered reliable. Most of the popular digital platforms were highly trusted on data safety. Digital platforms that scored high on trust quotient were also believed to obfuscate relevant information about the usage and storage of their data in lengthy terms and conditions. Four main reasons stated for not investing time to understand the use of software as service contracts were, There Is No Control (TINC), Fear of Missing (FOMO) on popular apps, there is no time (TINT) to read the lengthy terms and conditions and there is no option (TINO) as there was no alternative to the apps.

Participants outlined multiple instances where they saw advertising about topics that they had recently discussed in the vicinity of their smartphones. This led to a widespread belief that smartphone apps such as Facebook and Instagram were listening to their conversations. Industry participants were able to describe an extensive array of practices that were used to identify the right target audience for their campaign, using algorithms running on large databases built on behavioural and transactional data. These processes created highly accurate predictive abilities that enabled highly accurate profiling of Internet users. Some Everyday Users were aware of these processes and the analysis identified them as algorithm-aware. Others were not equipped or motivated to discover information to understand advertising technology. In the absence of this information, participants used heuristics to understand the ability of digital platforms to deliver advertising that is so relevant to their current personal situation. This common-sense explanation for personalised advertising (Ads are listening to me) is named Folk Theory of Customised Advertising. The research highlights the enabling nature of digital technologies in India and outlines a requirement for an easy-to-use toolkit for everyday smartphone users to become algorithm-aware and privacy conscious.

Table of Contents

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name: Ravi Dixit

Title of thesis: Download First, Privacy Later: Exploring the exchange of smartphone usage data for access to app-based services in India.

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission.

Signature: ..... Date: 21/12/2022

Acknowledgements

My research journey began six years ago in Gurgaon, India, carried on in New York City and culminated in Southampton UK. I am grateful to the New York Public Library, where I spent countless weekends pouring over its rich collection of books and journals via an effortless access system.

I remain indebted to late Prof Ashok Ranchhod who initiated me into this research, and for being my supervisor, mentor, friend and a guide until his very last days. I would like to thank Dr. Daniel Ashton whose sharp observations and patience kept me going for so long on this path, Dr. Joanne Turney who supervised me during a very important period and helped in my upgrade, and Dr. Yasmin Sekhon Dhillon for her guidance in ensuring that the end product is robust and well-timed. Thanks to the administrative team at Winchester School of Art who helped me with accessing the office, where I spent the last 12 months giving shape to this thesis after months were lost to the pandemic and related closures.

I have to specifically thank my awesome wife Vani for standing by me across three continents and constantly motivating me throughout my academic quest. I am thankful to my parents who inculcated the value of education and a zest for lifelong learning. I am happy that this tale of three continents has reached its finish line, and now begins a new chapter in my life.

Definitions and Abbreviations

Ad Tech ..... Advertising technology

AFDP ..... Advertiser funded digital platform

B2C..... Business to consumer

C2C..... Consumer to consumer

CRM ..... Customer relationship management

DSP..... Demand side platform. It is a marketing tool for automation of ad buying process.

EULA ..... End user license agreement

IDFA ..... Identifier for advertisers

IM..... Instant Messaging

PDS..... Personal Data Storage

PII..... Personally identifiable information

PPC..... Pay per click

RTB..... Real time bidding

SaaS ..... Software as a Service

SDK..... Software development kit

SERP ..... Search engine results page

SMU ..... Simultaneous media usage

SNS..... Social networking sites

TRP..... Television Rating Point

Chapter 1 Introduction

> What has been regarded as waste material – “data exhaust” – spewed into Google's servers during the combustive action of Search – was quickly reimagined as a critical element in the transformation of Google's search engine into a reflexive process of continuous learning and improvement. (Zuboff 2019, p. 68)

This chapter introduces my experience and expertise in the field of audience studies and a career path that led to this research initiative. The first half of this chapter outlines my interest in audience studies and establishes the aims of this research. Section 1.2 focuses on the growth of the media industry in India from a highly regulated industry to transforming into one of the most dynamic media markets in the world. Section 1.3 focuses on the business models of media in India and explains the development of advertiser supported mass media platforms. This section also documents the accelerated adoption of digital media platforms and the resulting explosion in data generated by the users of these platforms. Section 1.4 investigates the advances in audience measurement technologies and its impact on the growth of media platforms. It also investigates the impact of this growth on data experiences of the users. This section introduces concepts like the datafication of individuals and the role of advertising technology in the data economies of digital platforms. Section 1.5 describes the challenges for users in understanding the usage of their data and issues related to privacy in the digital world. The discussion about the research context in section 1.6 leads to framing the research questions.

1.1 Why research data experiences of everyday users in India

I started working in the Indian media industry a few years after CNN brought live coverage of the First Gulf War into living rooms across the country. Since Indian independence, terrestrial broadcasting was a monopoly of the public sector broadcaster and private broadcasters were not allowed to operate terrestrial channels in India. Though there were no regulations limiting the distribution of televisual content via cable networks as they did not utilize the airwaves. CNN and Star TV networks started beaming private TV channels into India via cable networks that were largely unregulated. Cable television was growing at a rapid pace when I started working with The Times of India in 1996 and started my journey into the field of audience research. After a stint in academia, I ventured into television audience measurement. As most privately owned television channels in India were heavily reliant on advertising revenues, healthy viewership ratings played a crucial role in their commercial success. The single most important factor defining the business success for responsible for business success for Indian TV channels was the ability to generate the

highest audience ratings and this was the common currency between media platforms, media agencies and marketers. Managing a television audience measurement organization equipped me to understand the construction of audiences and its role in media planning and marketing. Since then, understanding media audiences has been my lifelong passion.

I then moved to researching television audiences for The Walt Disney Company and then to market insights team of Google Inc and experienced the process of feature engineering where usage data was converted into audience cohorts to create customised advertising opportunities for marketers. Mass media in India like print, television and radio were always advertiser-funded and I had been part of sales teams that sold the "response" of audiences to advertisers even before the advent of mass accessibility of the internet. The key difference between traditional media and internet platforms was the method of measuring the audiences. While traditional media was measured through representative surveys for newspapers and longitudinal household panels for television, internet organizations were able to measure all their audiences to produce knowledge graphs unmatched by any other media format. Over the years, a host of digital platforms have developed non-advertising streams of revenue, but my interest has always been in understanding the audiences of media platforms and the processes of creating an audience commodity to be bought by marketers and sold by media platforms.

My work in the field of advertising effectiveness research for Google Inc helped me to obtain an under-the-hood look at the process of the creation of an affinity-based audience and design research to prove its superiority over broad demographic audiences of broadcast media. At a personal level I felt disconnected from the user experiences with algorithmic entities. This research enables me to document the processes for creating advertising audiences and the data experiences of people interacting with this process.

Throughout my professional career, I have studied audiences as a product to be sold to advertisers, research effort being solely focussed on improving the quality and quantity of audiences as desired. The key business objective of audience research was to demonstrate the role of mass media as an influencer throughout the purchase journey of a user.

This research is designed to explore the incentives of everyday audiences to indulge in the exchange of their data for access to advertiser funded digital platforms. This is driven by my curiosity to understand how audiences interact with the complex world of advertising technology.

There is a well-documented history of ad supported media reaping rewards based on their ability to measure their audiences. In India, audience classification has moved from simple age and gender demographics for traditional medium like print and television to complex affinity-based

audiences for digital platforms. Key selling proposition of digital platforms is their ability to deliver highly customised advertising to mass audiences to drive business results of the marketers. This chapter will provide the foundation of my research and explore the history of audience measurement in India as it grew from a single television channel country to one of the largest mobile internet markets in the world. The research will explore the understanding and awareness of customised advertising amongst the digital platform users in India.

As the name suggests, **Advertiser Funded Media Platforms** (AFMP) maintain a reciprocal relationship with their audiences and trade them with advertisers to fund services that are either subsidised or free to the end users. Digital devices have proliferated in every aspect of contemporary life with internet connected mobile phones being the most ubiquitous in India. These devices capture time-stamped data related to digital identity, geographic location, physical movement, attempted connection to Wi-Fi or cellular networks, usage of apps and storage of audio/video content. Individual platforms capture all the click-level usage data combined with device-level data and use feature engineering to create individual profiles that are used for classification, ranking and recommendations.

Digital platforms use this classification to get an understanding of consumer intent across various product categories. Brands in these product categories use these signals in combination with data from their Customer Relationship Management (CRM) databases and data brokers to create linkages across multiple platforms. These linkages like loyalty club cards allow advertisers and media agencies to link the activities in the online world with activities in the offline world. Physical location and offline purchases linked to digital usage are used to influence buying habits of the users.

There have been similarities between the rapid growth of cable television and the internet in India. In the initial phase both grew without any regulation, with no recognition of consumer rights and absence of a dedicated regulator. In the case of cable television, the users were left to negotiate service levels with cable operators that carried the channels to their home while and in the case of internet platforms there was no possibility of negotiation in accepting the terms and conditions which were a prerequisite to accessing the services. Over the years both the industries evolved, but digital platforms evolved in a more systematic manner as most of the popular digital platforms like Amazon, Facebook, Google, Twitter and Hotstar were owned by multinational corporations that implemented global systems in India. Revelations by Edward Snowdon and the Cambridge Analytica scandal brought the issues of data security and consumer privacy into the public domain. The aims of this research are drawn from my professional experience and interest

in the field of data experiences of everyday users of internet technologies with specific focus on customised advertising.

Aims

The research aims to:

- Understand the process of data generation by the usage of mobile devices and digital platforms and its use for audience profiling.
- Establish the level of understanding that Indian audiences have about the exchange of their digital usage data for access to app-based services.

1.2 Growth of Media

This section discusses the key components of the media business in India and the rapid rise of digital media platforms in India. As the media industry in India developed in spurts, usually driven by events outside the geographical boundaries of the country, instead of a smooth progression as witnessed in western countries. Introduction of colour television was necessitated by the request for a colour feed from other Asian countries when New Delhi hosted the Asian games in 1982.

The Gulf War of 1991 brought cable television to India and after that the number of private channels kept increasing. There was a resultant shift in advertising spends from print to television over the next two decades. CNN was the first cable channel to start beaming its content to India, followed by Star TV in 1991 and Zee TV in 1995 (Paul and Wang 1995).

Agrawal (2018, p. 13) described the liberalization of the Indian economy, resulting in advertising growth that funded the expansion of cable television in India. In the following decade, cable television increased its reach and was watched by 23% of the Indian population (Businessworld 2003, p. 158). This growth was achieved at the expense of print whose share of advertising fell from 70% in 1991 to 53% in 2001; the share of television advertising increased from 16% to 38% in the same period (Businessworld 2003, p. 182). In North America, print and television had a 45% and 37% share of total advertising and in Europe these shares were 54% and 33% respectively (Businessworld 2003, p. 182). In 2003, the share of advertising in Europe was strikingly similar to that in India.

The first phase in the transformation of media landscape in India was the growth of television reach and the next phase was growth in the ownership of mobile telephones. In the last three decades, India has gone from a country with low teledensity to being the largest smartphone market in the world (Bereau 2019). Ninan (2018, pp. 88–89) wrote about rapid growth in

teledensity in India coinciding with the economic liberalisation in India. In 1991, there were 0.6 telephone lines per 100 population and by 2002, the teledensity had increased to 4%. Ninan (2018, p. 88–89) wrote that the ensuing decade saw unprecedented growth as the total number of voice connections jumped to 805million by 2015. This growth in teledensity was riding on the rapid expansion of Indian economy that was growing by 7% since 1994. Agrawal (2018, pp. 3–4) predicted that by 2020, India's online community would grow to 700million amounting to 60% of the population. Purnell (2017) reported on the plans of Reliance Jio for providing 4G services to about 400million Indians who were using low-end phones.

In a short span of 12 years, the country has moved from 70% of the rural population being uncovered by multichannel television to being on the path to be covered by 4G internet. Ernst and Young (2019) reported that 96% of broadband subscribers were wireless users, consuming on an average of 8 GB of data per month. The adoption of digital technology also got a boost from the young demographics of India. It has the largest population of youth in the world. Spindle (2019) reported that half of India's 1.2 billion citizens are below the age of 25. Agrawal (2018) wrote about the transformation of Indian investments from Reliance Jio in rolling a nationwide 4G network, a 5-billion-dollar investment by Amazon and a rapid growth of Chinese smartphone companies fuelling the growth of mobile internet. These investments powered the rapid changes that took place in a short time period. Jain (2015) had attributed the slow adoption of mobile internet to high cost of 3G data. Comparing the outlook of Jain (2015) and Agrawal (2018) it is evident that in a few years India has jumped technologies and the reduction in the cost of internet data along with other factors powered India to becoming the home of one of the largest populations of smartphone owners in the world.

According to IMRB I Cube 2019 study, India had 627 million internet users that grew by 11% over the previous year. Urban India leads in internet penetration at 66% and rural India has an internet penetration of 25%. Internet access in India skipped fixed broadband and moved straight to mobile internet. According to a consumer barometer by Google (2017), 61% of internet users in India went online daily and 88% of internet users accessed the internet through their mobile phones. In comparison, 90% of internet users in the UK go online daily and 92% used mobile phones to access the internet. 78% of individuals in the USA use a smartphone and the number for the UK is 70% and 40% for India but most of the internet access in India is via mobile devices. An increased proliferation of mobile computing and mobile applications provides digital platforms with enormous amounts of user data. Goyal (2020) documented 25 years of the internet in India and noted "a massive 4.6 billion people today use the internet, sending out billions of gigabytes of data over the World Wide Web" and India had 600 million of these internet users.

1.3 Business of Media

Media platforms have two sets of customers, advertisers, and audiences. Media are in the business of marketing to their consumers as well as for their consumers. Warner, Lederer and Moroz (2020) describe the role of advertising in keeping platforms like Facebook and Google free for users and also outline zero incremental costs of inserting an advertisement into the content stream making advertising a highly profitable business model. Advertising income funds digital media platforms so they remain free to the end users and in turn marketers gain a mediated access to the users. In the early days of the internet, everything was free to access. Anderson (2009) described the expectation of "free" from digital platforms: "For the Google generation, the internet is the land of the free". Anderson (2009) described how advertisers cross-subsidized the users of free services like Google Search, Pandora, Skype, Facebook and MySpace. Since the early days of the internet, users expected online services to be free and were tolerant of advertising to keep it that way.

Audiences in India have been consuming advertising subsidised media products like newspapers and cable channels, in return for paying attention to advertising targeted to them. Digital platforms offered the audiences a similar exchange of their attention and usage data for free access to services.

As evident seen in the previous section, over the last decade digital platforms have permeated every aspect of life in India, enabled by mass availability of affordable mobile computing. Data management processes, regulatory frameworks and consumer data literacy are playing catch up with the accelerated adoption of digital platforms. In the advertiser-funded media ecosystem, access to information, surveillance and entertainment mediated via platforms is exchanged for attention to advertising message and usage data. According to Meeker (2019), there are 3.8 billion internet users across the globe and internet penetration has grown from 24% in 2009 to 51% in 2018. Meeker (2019) also reports that 53% of global internet users are from Asia Pacific and India has 12% of global internet users. In comparison, the USA has 8% of global internet users.

The top five technology companies in terms of capitalization (Microsoft, Amazon, Apple, Alphabet, Facebook) are headquartered in the USA. Alphabet and Facebook are two of the most valuable companies in the world and advertising revenue is the main source of their income. Domo (2020) reported that in an internet minute, 42 million messages are shared by WhatsApp users, 147,000 photos are shared by Facebook users, and 347,222 stories are posted by Instagram users. These figures illustrate the amount of digital data that is generated by half of the human

population while interacting with internet technologies. Indian users generate a large proportion of this data owing to their large share of the global internet userbase.

Zuboff (2019) described this data as the exhaust resulting from online activities of users and described how managing it has propelled Google to dominate the digital advertising business. Zuboff (2019, p. 75) called this excess data a "behavioural surplus" that constitutes the intellectual capital of Google and imparts a unique ability to predict user behaviour and sell advertising opportunities. These advertising opportunities as reported by McDonald and Clapp (2020) amount to $557.3 bn globally, out of which $330.9 billion is invested in digital platforms with mobile search attracting maximum spends. These figures show that digital advertising has overtaken all other mass media and one of the main reasons is their ability to target audiences with mass customization of advertising messages in real time. McDonald and Clapp (2020, p. 15) estimated that algorithms will be used to determine half the spend on digital platforms.

In the digital world, every activity generates a data trail and as data is generated from billions of activities from users, it is known as big data. According to McStay (2016, p. 135) big data "refers to an increase in volume of data, the velocity by which it moves and needs to be reacted to, and the variety of forms it comes in". Similarly, Kitchin (2014, p. 68) described key characteristic of big data as having a huge volume, high velocity, diversity in variety, is exhaustive in scope, fine grained in resolution, relational in nature and flexible and scalable. Big data by itself does not deliver value to the digital platform collecting it or the advertiser or users of these platforms.

1.3.1 Business of Big Data

Big data is not an exclusive domain of advertising technology and has existed from the beginning of computerised record keeping. Large datasets that can be indexed and accessed have existed since the advent of computerised data storage and retrieval systems. They were first created for stock markets, banking services, supermarket transactions and universal credit monitoring. Data about citizens has been collected and analysed by income tax authorities, welfare programs and law enforcement authorities. Marketing organizations in the consumer packaged goods (CPG) industry have been collecting data about consumer purchase habits via loyalty cards and media consumption habits via research agencies since computerised databases became commonplace. In other words, big data was getting collected long before the invention of the internet. Surveillance over people using data to understand their behaviour is also called "dataveillance". Clarke (1988, p. 499) described dataveillance as the use of decentralised databases that are connected via a communication network with cross identifiable records as a form of surveillance of an individual as they go through their day-to-day life.

Dataveillance in the digital sphere is possible because of digital traces captured and indexed by technology platforms. As discussed above, digital platforms were not the first custodians of big data, but they were the most motivated as their advertiser-funded business models depended on the ability to convert data into advertiser-relevant information. Mayer and Cukier (2014, p. 6) explain how big technology came to own the emerging field of big data.

> Internet companies could collect vast trove of data and has a burning financial incentive to make sense of them, they became the leading user of the latest processing technologies, superseding offline companies that had, in come case, decades more experience.

As most of the profitable digital platforms are funded by advertising, technology platforms had the right incentive to develop advertising signals from this data. The twin abilities of technology companies to collect and to create detailed profiles form big data have placed them in a dominant position as compared to other advertiser-funded mass media platforms. These detailed profiles are created by advertising technology that connects various parts of the internet for the purpose of monetization by digital platforms. Watkins (2019) described the various components of advertising technology (ad tech) as agencies, data brokers, services, exchanges and DSPs, advertising networks that sit between a marketer and digital platforms and facilitate the creation of custom audiences that enables the advertisers to target the right subset of users that are most likely to be impacted by advertising. Ad tech is the invisible layer of connectivity that links individuals across the internet and to the offline world. Ad tech also tracks an individual longitudinally and, for the purpose of advertising, throughout multiple purchase journeys that they are undertaking.

Einstein (2017, p. 149) characterises various stages in the consumer product purchase journey as “touch points" where digital marketers could interact with the consumer in the moments that matter. Einstein also explained how search history is used by marketers to target the consumers across the web with contextual ads. Einstein (2017, p. 149) outlined the process of matching the sales funnel of advertisers to the purchase journey and how consumers at the “need recognition” stage could be targeted with advertising to increase their brand awareness and consumers who are “evaluating alternatives” could be served with advertising to increase brand preference. This is one of the drivers for the growth of advertising on digital platforms. In addition to print and television, digital advertising has emerged as a major force in the last decade. Unlike the western countries, in India there is little discussion or debate about data privacy or any major initiatives towards imparting digital data literacy. The fact that India does not have a digital regulator or data privacy laws, makes it important to understand how the audiences are interacting with an

increasing complex advertising technology as described above. Historically it is similar to the initial years of cable TV industry where there was no regulation for distribution of television channels via cable.

This trend of subscription free access to news and entertainment (Print and Traditional TV) has migrated to digital platforms. Majority of mobile apps have a subscription free ad supported offering with a few offering a subscription offering that removes the ad. Hotstar from Disney, Liv from Sony, YouTube from Google are some of the examples of apps that have a subscription free advertising supported service.

Clapp (2018) brought out the unique nature of the media industry in India, as compared to the rest of the world; advertising on traditional media like print and television is still growing here albeit at a slower rate as compared to digital advertising. Data from WARC (2019) showed that television has 35%, print has 34% and the internet has 27% of the share of total advertising in India. These figures make India one of the few countries with a flourishing print media.

According to Ciochetto (2013, p. 67), newspapers and advertising practices were first introduced to India by the British and even today the practice in India is influenced by global conglomerates. Over the years, media planning in India has largely followed the practices laid down in the Western world. According to Ernst and Young (2019, p. 233), nine out of the top ten media agencies in India are part of global organizations. According to Nordenstreng and Thussu (2015, p. 105), "The BRICS countries account for 40 per cent of the global population and a quarter of the global GDP, the study of media in these nations remains largely neglected in international media studies" and this neglect of studies in the practices of media organizations in India is also reflected in the lack of studies looking at the awareness of data practices of digital platforms amongst Indian internet users. According to Kennedy (2018, p. 18), "Despite the significance of such everyday practices in the production of large-scale data, little attention has been paid to people's thoughts and feelings about these data-producing processes." Lupton (2018, p. 3) categorizes the study of people interacting with a complex technology environment via user interfaces: "Data sense, therefore, may be conceptualised as the co-constitution of human and nonhuman sense-making." My research aims to address the lack of research into data sense and how customised advertising is experienced by Indian smartphone users

This section establishes the rapid growth in the adoption of mobile internet and advertiser-funded business models of some of the largest digital platforms. A historical lack of regulation for new technologies has resulted in a situation where Indian users do not enjoy a regulatory oversight over their relationship with digital platforms. A large portion of the population has jumped from basic voice-only phones to smartphones in the last five years. This presents an

adequate pool of subjects that can be researched. This transformation of India from a country with low teledensity to one of the largest smartphone markets in the world presents an opportunity to explore data awareness and attitudes towards advertising technology amongst a population that is at the forefront of mobile internet adoption. My research stems from this rapid growth in the adoption of mobile apps in India and a lack of India-specific research in the field of media studies (Park and Curran 2000, Nordenstreng and Thussu 2015). A critical component of creating advertiser friendly audience is the ability of media platforms to measure their audience and report the advertising exposure of advertising campaigns amongst the marketers target audience.

1.4 Evolution of Media Audience Measurement

Media ecosystems developed slowly after the advent of the printing press and it was not until the arrival of radio and television that change gathered pace. Radio and television both developed in terms of hours of programming and geographical coverage powered by advertising revenues. Marshall and Quentin (1967, p. 14) called radio and television “electric media” that created a “worldpool of information”. With the advent of electric media came the ability to simultaneously reach large audiences across wide geographies. Since the very early stages of mass media advertising, both media owners and advertisers were interested in understanding the impact of advertising and content. In the United States, advertising-funded radio started early and better measurement improved its profitability. Webster et al (2013, p. 21) outlined the increase in the profitability of radio between 1930 to 1935 and television stations after 1950. This growth was triggered by better audience measurement. Webster et al (2013, p. 173) provide the rationale for television audience measurement as “the need of advertisers to buy audiences, and the eagerness of broadcasters to sell them, brought the ratings service into being”. Television ratings providing viewership, measured through nationally representative samples, gained the status of currency that was used to trade audiences. These viewership and readership numbers were used to create media plans to deliver the right message to the desired audiences. This illustrates the importance of measuring and classifying audiences for AFMP.

The starting point of any communication plan is identifying an audience segment that is receptive to the brand and a media plan is the tool to deliver this message through a selection of media platforms. In describing market segmentation, Reynolds (2006, p. 445) explained that “the basic objective of this function is to identify homogeneous groups or customer segments in the marketplace that will respond in a consistent, predictable way to variations in the marketing mix”. As discussed above, until the late 1950s, audiences had been measured, reported and planned on the basis of demographics. In the USA, there was direct competition between newspapers and

television channels to compare their demographic audiences. According to Warner et al (2020), age and gender were the two most commonly used segmentations. As a number of television channels started multiplying so did the complexity of the audiences measured.

Similarly in India, audience measurement developed with the rise of multi-channel television. According to Khandekar (2013) there was only one national terrestrial channel in India until 1991, television audience measurement systems did not have to grapple with the same issues as those in the USA. In the next two decades the number of television channels multiplied,

driven by rapid growth in GDP. Television industry kept pace with economic liberalization and the number of TV channels went up to 450 (Athique 2012). At this point the challenges of media fragmentation in India were matching up to the complexity of Western countries.

Traditional representative audience measurement panels suffered from sample size issues and were unable to provide detailed demographics beyond age, gender and geography. Green (2017) listed out five stages of audience measurement (AM) research in the USA. It started with counting units or copies of publications sold (AM 1.0, 1914), counting people or audiences for publications and radio (AM 2.0, 1939), broader and deeper measurement with the introduction of peoplemeters in 1949 (AM 3.0) that were the mainstay of data until the early 80's. This pattern was broadly replicated across the globe with different adoption timelines. AM 4.0 was envisaged when television audiences started to fragment in 1980s. Green (2017) described AM 4.0 as the beginning of online audience measurement where the focus was beginning to shift towards cross-media measurement. Green (2017) wrote about the challenges of audience measurement for digital platforms and imagined that AM 5.0 was where everyone was struggling due to multi-screening or simultaneous media usage. Multi-screening meant that the same minute can be attributed to multiple platforms, leading to the challenge of time fragmentation. None of the measurement systems that were designed in phases 1 to 4 were capable of handling time fragmentation. India reached the complexity of AM 4.0 around 2012, as channels multiplies across genres, languages and geographical areas (Athique 2012).

Advertisers buy media for the eyeballs that are glued to the screen. Audiences started fragmenting since the 1980s with the arrival of multiple channels in North America and in India post-1991. With the advent of the internet and mobile internet, the number of screens multiplied, and the attention got divided across screens. Wolf (1999) wrote about attention being a precious commodity and a growing time poverty of consumers. Wolf (1999, p. 35). summed it up as "consumers today will happily sacrifice money to gain time" and what it meant for advertisers was to imagine a way to provide a relevant experience to engage the audience and have an impact. Pine and Gilmore (1998, p. 98) named such experiences as "a memorable event" that

brought goods and service together in what was described as an “experience economy”. Another important concept for television viewership was the concentration of viewership in a narrow time band called “prime time”, where the highest concentration of audiences would converge. Wu (2017, p. 94) defined prime time as “the attentional habit of turning on the radio (later, the television) at the designated hour each and every evening of the year”. This successful formula was successfully applied across the globe. The television prime time was based on the data for audiences gathering around radio or television at the time when compelling content was aired. Advertising in the age of prime-time television was about aggregating the largest possible audiences and broadcasting an advertising message to them. Wu (2017, p. 263) described television advertisers as “the attention merchants” bombarding audiences with legally permitted advertising. Another concept that is related to audience fragmentation was attention fragmentation that emerged with the adoption of the internet and increased occurrences of simultaneous media usage (SMU).

Mobile internet and apps accelerated both time fragmentation and attention fragmentation due to SMU or multi-tasking in front of the television. Digital advertising technology powered by big data analytics brought a detailed knowledge of audiences to the marketing departments and developed a system whereby relevant advertising was served when the customer was primed for a purchase. Ad tech kept evolving and was able to provide advertisers with the ability to target audiences based on their affinity to a particular product category or interest in purchasing a particular brand. According to Levy (2011), Google was a very successful consumer entity with its search engine but the AdWords platform that delivered textual advertising made it into a technology powerhouse that was profitable and cash rich. Digital technologies have brought an end to television prime time as the only premium media real estate. Due to rise in multi screening, an always-on mobile screen is considered to be the new prime time. Research done by Dixit (2016) on the impact of synergy between television and mobile video advertising showed that in India, 90% of users who access the internet on their smartphones consider it their primary device. Batra (2017, p. 185) showed that time spent on a mobile device in a day is more than the total time spent on television in a week. As consumer media consumption habits are shifting, advertising spends are following suit, but need better audience measurement systems to manage this changing media ecosystem. One solution is for audience measurement to upgrade to AM 5.0 and to measure audiences in a postdemographic media trading system. Apart from third-party audience measurement, marketers are also asking for more data about the audience of digital platforms to improve the efficiency of advertising spends.

In a massive change from the audience estimates of broadcast media generated from sample-based panels, digital platform audiences were not mere demographics that could be accurately

measured by a small representative panel. Elmer (2003) described the defining characteristics of digital platforms as their ability to collect data about their consumer universe and converting it into advertiser-friendly targeting cohorts. McStay (2018, p. 97) further elaborated on the process of converting people into "data doubles" and mapping this virtual identity to "interests, outlook, beliefs, emotions, age, life-stage, income level, relationship status". Individual users of digital platforms were being converted into affinities and intents that could be used for advertising products and services. Szulc (2018, p. 3) described "datafication"; the practice of creating profiles on social networking services (SNS). Datafication of audiences provided a key differentiation to digital platforms as compared to broadcasting platforms.

1.4.1 Emergence of Postdemographic Audience

Evolving media economics resulted in a rush for aggregating audiences and advertiser funded media in the USA and India. Wolf (2017) described this as the opportunity that digital media seized on with a promise of "absolute measurement". At the core of this measurement was audience data that traditional media could never generate about their audiences. There was an increased requirement for sorting and profiling digital media audiences as advertising started shifting to digital platforms. On the other hand, marketing organizations are also collecting an ever-growing volume and variety of data about their customers. According to Einstein (2017, p. 159) when audiences are using a free digital service, there are "multiple advertisers bidding at auction for attention" and all these data sources come together via ad tech to create topical communications opportunities. This datafication of digital audiences is also called a postdemographic audience.

Rogers (2009, p. 1) defined postdemographics as "the study of the data in social networking platforms and in particular how profiling is or may be performed". Today, it is possible to extend this definition of postdemographics to include the practice of big data enabled profiling to all the digital platforms. Straubhaar et al (2013) presented an audience-centric model of media consumption where all media converges and a number of interactions between audiences and media format generated an electronic trail of data that powered the creation of postdemographic audiences. Grimshaw (2017) in an introduction to the Admap issue labelled "The Death of Demographics – How to Define and Reach the Audience", wrote that "the obsession with an age demographic is all the more curious when the traditional broad demographic ways of segmenting an audience have been consigned to the dustbin for many/most marketing categories". This issue of Admap outlined the upheaval in the marketing communication industry as it came to terms with the complexity of segmenting and targeting consumers in the digital advertising ecosystem.

Davisson and Booth (2017, p. 16) described the process by which Facebook generated revenue by selling user data to advertisers and this data includes information about age, gender, location, friends or social connections, interests, hobbies and visits to other sites and apps. Davisson and Booth (2017, p. 16) brought up the difference between broadcast media usage where the audience traded “viewing of advertisements” for access to content, with Facebook audiences “laboring further to create entertainment content”. This usage data is further utilised by platforms and advertisers for better product experience and targeted advertising. As discussed above, rapid growth in digital platforms ensured that broadcast media audiences that were essentially demographic in nature were competing with detailed profiles of digital audiences. These profiles are created through data practices that collected and analysed a much wider set of variables about individuals, something that was not possible to be collected for broadcast media audiences. Marketers still needed more data and insights from the digital platforms to feel comfortable shifting their advertising spends from broadcast to digital platforms.

The process of converting this data into marketing tools to understand the efficiency of advertising keeps improving and these changes indicate an ever-growing desire of marketers to demand data that can help them target audiences in a cost-efficient manner. Nielsen released its first Chief Marketing Officers (CMO) report in 2018 and only 26% of the CMOs were highly confident about accurately quantifying their digital media return on investment (ROI). The fact that there is a feeling of data deficiency at the very highest level of a marketing organization is an indicator that they need more data signals from digital platforms. On the one hand there is an increasing concern about the amount of data that is collected by digital platforms and on the other hand there is a feeling that not enough of this data is being shared with advertisers. This research will focus on exploring the practices of data usage by advertisers in India to understand how data privacy is built into the marketing process.

Despite this lack of tools to measure their digital ROI, 80% of the CMOs expected to increase their digital spends in the next 12 months. Television was still the most important medium for half of the CMOs. 23% of CMOs were highly confident in accurately quantifying the ROI for their traditional media spends. Even after two decades, a majority of CMOs were still not highly confident of measuring the return on their advertising spends in the digital medium. At the same time, digital media is increasingly gaining in its reach, usage and importance. The increasing demand for audience data is evident in Nielsen's CMO report, however, professionals outline the lack of regulation around data practices of digital platforms and data mining organizations (Kennedy 2016). Prior to the enactment of GDPR in Europe there was a felt need for an industry regulator as there were fly by night operators in the field of social data mining who had scant respect for consumer data privacy. Such operators were described as "cowboys" by a research

participant in Kennedy (2016, p. 123) and even other respondents in that study cited lack of rules surrounding the demarcation between public and private data.

This demand for more data and accountability from digital platforms raises the question of data privacy. There has been a slow reaction from regulators to protect consumer data from being bought and sold as reported by Klosowski (2021) and Apple seized this opportunity to make consumer data privacy its key message. Haggin and Vranica (2021) reported on the privacy-centric features in iOS 15 that will restrict access to a variety of identifiers like IP addresses to third-party data brokers and would also put brakes on email tracking. The next section examines the history of predictive algorithms and issues surrounding data privacy.

1.5 Data and Privacy

As algorithms have improved their abilities to predict user behaviour, they also bring up the issue of data- and algorithm-awareness amongst everyday users of these digital platforms. Looking back a decade, Thompson (2010) reported Schmidt discussing the future of Google Search:

> At the same time, Schmidt envisions a future where we embrace a larger role for machines and technology. “With your permission you give us more information about you, about your friends, and we can improve the quality of our searches,” he said. “We don't need you to type at all. We know where you are. We know where you've been. We can more or less now what you're thinking about.”

Eric Schmidt was discussing user permissions and the role of user consent in collecting metadata that provides context to a search query. Eric Schmidt was also describing the future of Google Search where the questions would be predicted, and search results would be shown before the question is asked. In 2010, Google was preparing for a future to predict user information and entertainment needs and answer questions that have not yet been asked of its search engine.

Google Search and its predictive abilities to provide answers is one such example. Auletta (2018, p. 167) compared the attitudes of youth regarding privacy in China and the USA and concluded that they were pretty much the same. Auletta (2018, p. 167) outlined the informed exchange that youth in both the countries were having with the software service providers: “Younger generation maybe willing to trade their personal data, perhaps for discounts or partially subsidised content”.

At times, users find advertising as useful information and the data that is generated by accessing free services is considered by McStay (2018, pp. 97–98) as “meaningful information used to affect behaviour and experience”. Ad tech and related ad serving engines use behavioural and transactional data to create profiles that are used to segment and target audiences for advertising

campaigns that seek a positive behaviour change in the favour of the advertised product or service. The main method of creating these audiences is to develop the ability to track them across the internet and the offline world. McStay (2017) states that digital platforms manipulate this data by reducing target audiences to “cookies and device IDs” amongst other identifiers for advertising.

While predictive power of algorithms has kept up with the predictions from a decade ago, consumer's ability to control the data that enables these predictions has failed to keep pace. According to the National Privacy Test Report of Nord VPN (2020), Indians scored the worst in the test and had the lowest score for any country in reading the terms and conditions for digital platforms. The report outlines a global trend where users do not pay attention to the End User License Agreement (EULA) that they sign, but Indian users performed towards the worst.

Apart from the issues with explicit informed consent, there are other privacy issues where the vulnerability of mobile operating systems enables passive surveillance that is beyond the control of everyday users. Some of these privacy challenges became evident in investigations by Tau (2020) wherein a company named SignalFrame was able to track people using Bluetooth signals on Android devices. In another Wall Street Journal investigation Schechner et al (2020) looked into location tracking practices of a data broker X Mode Social where it was discovered that it was combining location data across multiple sites and apps and sharing it with federal authorities; Foursquare Labs Inc and Cuebiq Inc were also found to be indulging in such practices. These investigations bring the issue of moral limits to the use of technology and aiding governments to keep tabs on its citizens. The two examples of The Wall Street Journal investigations show the privacy boundaries that are being breached regularly and the rise of dataveillance in unimaginable ways.

An increasingly complex nature of data collection as discovered in The Wall Street Journal's investigation above, makes consumer privacy and the use of technology for surveillance an important area of research. Overall, the ubiquitous nature of mobile internet and internet-enabled connected devices, consumer privacy, informed consent, and data experiences form an important element of the research context.

1.6 Research Context

Previous sections have established the mass adoption of digital platforms and the resultant data traces that power the revenue generation of these platforms. Concepts like datafication and lifelogging have been introduced and journalistic investigation and scholarship has been referenced to highlight the privacy implications of increased dataveillance. This section

establishes the context of this research and introduces the importance of the voice of the consumer in the formulation of strategies to mitigate the impact of dataveillance and legislative initiatives for the purpose of oversight.

This study explores everyday data experience of individuals and their understanding of algorithms that shape these experiences. Customised advertising is one such case of interactions of users with ad tech. Ruckenstein and Granroth (2020) describe the frustration experienced by their participants when the ad tech fails to serve relevant advertising but are thrilled to receive recommendations and relevant advertising. Lomborg and Kapsch (2020) outline the varied nature of algorithmic awareness amongst their participants and categorised three sources of knowledge: professional knowledge, experience based knowledge and third party knowledge. Lomborg and Kapsch (2020) found that participants deployed multiple strategies when faced with algorithmic entities ranging from adulation to activism.

Successful digital media platforms collect large amounts of data to enhance the quality of services for the users. A better platform experience drives more usage and thus generates more data creating a virtuous cycle. Swisher (2019) called user data “the rocket fuel of the ever expanding internet” and as seen above the rocket fuel has powered a rapid growth in the adoption of digital platforms across the globe with India leading the world in the adoption of mobile internet.

Glass and Callahan (2014, p. 43) called Google and Amazon a “data muscle” where more usage of their services made their muscle stronger. They used multiple examples of companies like Dell and Sears to showcase the use of customer data to build strong businesses in the pre-smartphone era. Glass and Callahan outlined that the practice of organizations developing data muscles in the pre-internet era and the collection and analysis of big data was heralded by the mass availability of computing and digital platforms improved upon this practice and grew muscles that were stronger than any other organization in the history of marketing. Glass and Callahan (2014, pp. 135–136) talked about how “Netflix flexes its data muscle” by building connections between different pieces of content watched by the same users. Netflix used the power of its data muscle to cross-promote its shows to audiences watching the same artist or genre of storytelling. Shoenberger (2017, pp. 300–301) discussed the benefits of digital advertisers churning huge amount of user data to create relevant advertising but raised questions about the issue of consumer knowledge and control over their usage data.

According to Facebook (2019), audiences can be segmented and targeted across multiple dimensions like 1) demographics, 2) interests, 3) location, 4) behaviour, 5) custom lookalike audiences and 6) brand site visitors/brand app visitors. In the digital media ecosystem dominated by mobile internet access, information flow has changed from unidirectional to bidirectional.

There are numerous examples of the collection and sorting of consumer data and thus it is important to understand the awareness and attitude of users of digital platforms towards the underlying advertising technology.

The mass adoption of mobile internet has brought increased challenges to personal privacy, in terms of customer datasets as well as audience measurement. Vega (2012) reported results from a survey where a majority of voters expressed disinterest in political ads being tailored to their interest. The report also mentions the request of the FTC to Congress to enact legislation to protect the privacy of internet users. In an editorial published on the 19th March, 2011 the editorial board of The New York Times was demanding an internet privacy law. Kandias et al (2013) scrapped publicly available data about YouTube videos and were able to determine the political views of people who commented on those videos. On the other hand, privacy concerns have not prompted mass movements away from digital platforms despite revelations from Edward Snowdon and the data scandal involving Cambridge Analytica (Kantchev and Gross 2018). Privacy concerns have forced digital platforms to increasingly limit access to consumer data to third-party measurement providers. This proved to be the biggest challenge for researchers to understand digital audiences and for advertisers to know the demographics of consumers that buy their products as well as consume digital content.

A case in point is one of the largest CPG advertisers, Proctor and Gamble (P&G). Vranica (2018a) reported that P&G's pushback began more than two years ago. Vranica (2018b) reported that Unilever threatened to pull advertising dollars from digital platforms citing brand safety issues and stressed that advertising spending billions of dollars should speak out about issues on tech platforms. This is another example of advertiser frustration. Bruell (2017) reported another instance of Marc Pritchard warning that tech companies are not doing enough to open themselves up for third-party measurement of their audience claims as well as the viewability of ads of digital platforms.

There is no data privacy law that protects the users of digital platforms, and the USA does not have a federal law enabling users to access or control their data. Klosowski (2021) reviewed various privacy laws across the USA and found that only three states had laws that enabled the residents to exercise control over their data and its use: “California (CCPA and its amendment, CPRA), Virginia (VCDPA), and Colorado (ColoPA). Regardless of which state a company is located in, the rights the laws provide apply only to people who live in these states.” According to Klosowski (2021), a proposed law aiming to increase privacy protections for the citizens failed to pass in North Dakota amid a strong opposition from some most prestigious organizations representing advertisers, digital platforms and media agencies: the Association of National

Advertisers (ANA), the American Association of Advertising Agencies, the Interactive Advertising Bureau, the Network Advertising Initiative (NAI) and the American Advertising Federation (AAF).

There are opposing forces being applied by advertisers on digital platforms to share more data to improve the efficiency of media spends and a countervailing drive for improving privacy protection through legislative means places the voice of the consumer at the centre of this debate. There is scholarship that makes a compelling case for a change to "planetary infrastructure of surveillance" as theorised by Sadowski (2020, p. 173) and an equally compelling case for including "everyday experiences" of people informing data activism Kennedy (2018, p. 18). Papacharissi (2010, p. 306) emphasises the importance of individual and their interface with technology: "the architectural environment presented through these affordances places the individual as the centre and source of all interactions."

To summarise, scholars (Papacharissi 2010, Kandias et al 2013, Glass and Callahan 2014, Iliadis and Russo 2016, Kennedy 2018, Lomborg and Kapsch 2020, Ruckenstein and Granroth 2020, Sadowski 2020) guide the focus of this research on individual users of digital platforms. The research questions are aimed to contribute to the scholarship with a focus on everyday experiences of young Indian users.

1.6.1 Research Questions

All mass media formats in India are largely funded by advertising revenues. In the era of print and television, a cohort of audience were traded with advertisers and generally the audiences were unaware of this exchange. This trade has become complicated in the digital world where both audience attention and their usage data are used to create audiences for advertisers. Scholars like Watkins (2019) Einstein (2017) McStay (2016) have described the advertising technology that generates majority of revenues for digital platforms. There is a general lack of research in the field of everyday data experiences of Indian internet users. My aim is to understand the interactions between users and a highly automated, dynamic advertising targeting system that generates customised advertising informed by their "data exhaust".

All these factors bring us to the following two questions that this research addresses:

Q1. How do young Indian smartphone users experience personal data and its reciprocal use by digital platforms to provide access to software services? This question is focussed on:

- Individuals aged 18 to 34, residing in top two metros of India and regular users of Amazon, Facebook, Google Search, Twitter and YouTube.

Q2. How are audiences for customised advertising created from digital data trails of users? This question is focussed on:

- Practitioners working in Mumbai and Delhi in the field of marketing.

This chapter describes the role of audiences in fuelling the big data systems that drive the ability of ad tech to predict future behaviour and serve these affinity-based audiences to advertisers. This rapid adoption of technology has occurred with a regulatory oversight in India.

1.7 Thesis Outline

Chapter 1 establishes the aims of this research leading to research questions. Chapter 2 examines scholarship across three main themes of data experiences, data privacy and surveillance. The review developed a perspective on the deployment of advertising technology to create communication opportunities for advertisers. Chapter 2 investigates the concerns of scholars about interconnected databases, datafication of individuals, profiling, algorithmic sorting, and the creation of dataveillance tools. This chapter tracks the path from the audience commodity of broadcast media to the postdemographics commodity of digital media. The creation of an audience commodity and its evolution into a powerful targeting tool by digital platforms and programmatic ad serving mechanisms was the key outcome of the first half.

The review of privacy leads to an exploration of developments in scholarship on the use of Bentham's panopticon to surface issues with advertising surveillance and the exchange of privacy for convenience. The review then returns to audience profiling and datafication of users and compares it with the prevalence of participatory surveillance. Increasing complexity of data gathering practices of AFDP in India presents an area of academic research that has not been explored in detail. A need emerges for understanding data experiences of everyday users of digital platforms and analysing the role of advertising technology on evolution of common-sense theories. These discussions highlight that there has been a lack of focus on research examining the understanding of consumers related to advertising technology. The literature review identifies a lack of scholarship about research about Indian audiences and their data experiences with advertising technology.

Chapter 3 introduces the research design and maps suitable research methods to answer the research questions. Qualitative research methods and personal and group interviews were used to evaluate contemporary experiences of participants engaging with their personal data. This research instrument was designed to explore the process of converting the digital data trail into customised advertising for smartphone owners and their incentives to participate in this exchange

of data for access to app-based services. It goes on to describe the sampling approach and the recruitment process. The next half of the chapter describes the process of data generation, the research instrument, and its deployment in the field research. The last section describes the approach taken to analyse the data through data flow diagrams that bring out the richness of data and reoccurring themes across interviews.

Chapter 4 draws a comparison between scholarship that discusses various practices of data collection, how aggregation of data across multiple owners creates privacy risks for the everyday user, and the extent to which this was understood by the participants. Further it explores the role of big data and the participants' understanding of predictive algorithms that powered product features to improve adoption and engagement with various digital platforms. The chapter explores the evolving nature of data exchange between participants and digital platforms as well as between digital platforms and advertisers. The discussion moves to how this dyad of exchanges has been replaced with a network of interactions, the inherent invisibility of these networks and lack of control experienced by the participants. This chapter analyses key mitigation tactics of consumers when faced with complex EULAs and the invisible interconnectedness of different aspects of advertising technology. These data experiences are contrasted with the literature on media studies from other countries along with scholarship detailing the ad tech components of identifying and tracking individuals across the internet. The key concepts that emerge are, components of trust threshold and tactics used by participants to mitigate against data disquiet.

Chapter 5 explores the contours of privacy and describes the privacy expectations of the participants with regards to the use of their data on digital platforms. Scholarship on privacy paradox is contrasted with the findings of my research and participants reflections on the exchange of privacy for convenience is recorded. Deployment of a new research methodology called privacy audit is outlined where the participants were encouraged to discover their own digital profiles stored in the mobile apps, and to reflect on the data practices of digital platforms. The second half of this chapter highlights the privacy concerns of participants and lists out redlines that they do not want to cross. Privacy concerns expressed by the participants are compared to those expressed in the literature and the role of trust threshold in navigating privacy redlines is also described. This section also reflects upon the social aspect of algorithmic recommendations in the Indian context. Industry Experts describe the state of privacy discussions in Indian organizations while executing online advertising campaigns. This chapter culminates with Industry Experts detailing the practices and processes that result in creation of customised audiences for advertisers and inner workings of ad tech that power these recommendations.

In Chapter 6 Industry Experts describe the sources of data available for creating qualified audiences from the general population using digital trail data and the capabilities to deliver customised advertising to this audience. Industry Experts described the use of DMPs for enriching the profile of consumers. They went on to define the first-, second- and third-party databases and the ability of ad tech to identify consumer journeys and track cohorts across the internet. A detailed list of identifiers that are used to track individuals is documented. Everyday users describe their encounters with algorithmic profiling, automated targeting, personalised recommendations, and customised advertising. There is an in-depth exploration of how these encounters shape the attitudes of participants towards personalised advertising. Overall, this chapter maps various types of surveillance described in the literature and compares the views of scholars about advertising surveillance with that of participants. The chapter concludes by contrasting scholarship about common-sense theories developed by everyday individuals about their interface with a complex information environment and compares it with folk theories developed by the participants of my research with regards to personalised advertising and other data experiences.

Chapter 7 provides a summary of the research along the three themes of Data, Privacy and Surveillance. It goes on to highlight the 5 main findings of my research, its implications, and areas for future research.

Chapter 2 Literature Review

2.1 Introduction

The previous chapter introduced three themes that guide the review of literature, and in this chapter, scholarship is reviewed across these themes of data, privacy and surveillance. Generally, this review focuses on literature published until 2019. As CDS is a nascent field some of the more recent literature is also reviewed to keep the discussions topical.

Section 2.2reviews scholarship related to datafication of individuals, use of code for dataveillance and relevance of panoptic sorting in contemporary literature about predictive data practices. This section focuses on scholarship on the concept of postdemographics and its implications on privacy. The section also explores data practices of digital media platforms from the early days of the internet to the era of mobile computing and the impact of such practices on users. The review examines literature on the accelerated adoption of digital platforms and the resulting data explosion and follows the developments in the field of audience commodity and changes in this domain. Section 2.3 discusses various definitions of privacy and their applicability in today's data rich environment and varied perspectives on the legitimacy of privacy practices. The review investigates literature about the state of consumer data privacy, the impact of datafication on the privacy of individuals and compares mass personalisation of recommendations heralded by digital platforms and its impact on data privacy. This section also explores literature about privacy concerns of individuals in the digital world where echoes of digital behaviour do not expire. Section 2.4 follows up from the previous section and explores the literature about technology enabled surveillance and will focus on the scholarship around advertising surveillance. This section maps literature on participatory surveillance to the practice of Quantified Self. Section 2.4.5 discusses the use of heuristics by everyday users of digital platforms to create an understanding of the complex technological environment that enables personalised advertising. The main focus of this section is folk theories related to algorithmic recommendations. Section 2.5 enumerates the research gap that this research seeks to address.

Jenkins (2006) described various facets of contemporary media landscape as "innovative", "convergent", "everyday", "appropriative", "networked", "global", "generational" and "unequal". These descriptions convey a multifaceted impact that the convergent internet technologies have had on the media industry and its users. An accelerated adoption of internet technology has created a network of devices as well as individuals and also created a digital divide between digital haves and have-nots. These technologies fundamentally altered the structure of mass

communication and the Source-Message-Channel-Receiver (SMCR) model by Straubhaar et al (2013, p. 18) became increasingly redundant. The linear SMCR model was developed to explain the largely unidirectional flow of information and entertainment with no return path for the data from audiences to flow back to the media companies. The model was well suited when media companies were owned by larger complex organizations that controlled the availability and distribution of information and entertainment. Castells (2013, p. 55) compared "one directional" mass communication of the past, to the "mass self-communication" of the digital age. One of the first genres to be disrupted by this transformation was news.

Given these massive changes, the concept of news and its ownership by complex media organizations was challenged. With social media platforms and instant messaging (IM) a story can be amplified at an unprecedented pace. Castells (2013, p. 121) wrote that "the evolution of news towards infotainment broadens the scope of consumerism to the entire social and political realm". News media that used to set the agenda for public discourse is now reporting on the agenda set by public discourse occurring in the online world. This was made possible with citizens organizing on digital media platforms and a growing mistrust of the expert view of traditional gatekeepers. In the digital media ecosystem dominated by mobile internet access, information flow has changed from being unidirectional to bidirectional.

In the hyper-connected world (see section 1.2) of IM, social networks and mobile access to audio–visual stimulus, there is an abundance of avenues for getting information. Audiences acquire their understanding of the world around them via traditional and digital media platforms. Protess and McCombs (2016, p. 2) stated that "Individuals construct their own set of pictures from their direct personal experience, what they read in books, magazines and newspapers, and what they see on television and in movies". As the audience shifted their source of information and entertainment towards more interactive digital formats, so did advertising and they followed them on to the internet. Digital audience-targeting capabilities allow advertisers to customise communication on the basis of editorial context and audience affinities.

Einstein (2017, pp. 164–64) describes the collection of big data by digital platforms and its use by marketers to determine the position of their target audience in the "buying process" and to predict the type of products that the audience could be interested in. Personalization of communication presents opportunities to target audiences' needs with personalised creatives. This is very different from advertising in a traditional mass media where a similar set of creatives are served at scale to an undifferentiated audience. Young (2017, p. 21) describes the impact of digital media on advertising agencies as "explosive" and points out that not only has the digital age brought about "audience fragmentation" but also "disintermediation". This disintermediation

from broadcast media meant that the internet took over the role of mediating the relationship between audience and content, as well as advertisers and audiences.

2.2 Data

There is scholarship that investigated use of large-scale computerised databases to profile and sort users. Bogard (1996) viewed profiling as the first step towards building a surveillance system and nudging the profiled population towards a desired behaviour. This is similar to how digital platforms use digital data trails to build profiles and nudge the users towards the goals of engagement on their platforms and also serve as mediators between the advertisers and users. Bogard (1996, p. 27) described "computer profiling" as preparing the groundwork for a surveillance system to function effectively. He also called it "observation before the fact". These types of profiles need to be constructed ahead of time as most of the bidding in advertising technology happens in real time. In the digital world, information derived out of profiling and used to predict future behaviour by using past data and profiles is the product offered to advertisers.

Elmer (2003, p. 9) added more dimensions to profiling by digital platforms and described it as "a process that focusses on the collection, storage, networking, diagnosis, and deployment of demographic and psychographic information". As discussed in section 1.3, advertising revenues are the mainstay of any profitable media platform in India. The relationship between the advertiser and the media platform is evaluated on the ability of a media platform to deliver target audiences for advertising campaigns. Advertising technology that creates and delivers qualified audiences for advertisers is built on computerised profiling as described by Bogard (1996) and along the dimensions listed by Elmer (2003). Einstein (2017, pp. 164–64) described the ownership of big data by digital platforms and its use by marketers to determine the position of their target audience in the "buying process". Profiling or sorting of the audience into interest categories and the ability of ad tech to deliver customised advertising has created some of the largest corporate entities in the world (see section 1.3).

This digital existence of users is held together and directed by software code. Berry (2011, p. 33) explains the code as both force and a boundary parameter: "Code is striking in its ability to act as both an actor performing actions upon data, and as a vessel, holding data within its boundaries". Profitable advertising-funded digital platforms listed above are dependent on increasing user interaction and advertising that results from it. Code is used as a means of achieving the desired results; for users it provides access via a user interface and for digital platforms it drives business outcomes. The role of code is directly related to the business model of digital platforms. As

described by Van Dijck (2013, p. 170), “free” access to digital platforms means that some other entity paid for the access in return for user attention and their transactional and behavioural data.

Berry (2011, p. 145) described the changing nature of the internet from a data indexing and retrieval system to an amalgamation of constantly evolving flows where the users are both the receivers and generators of data: “The user becomes a source of data too, essentially a real-time stream themselves, feeding their own narrative data stream into the cloud, which is itself analysed, aggregated, and fed back to the user and other users as patterns of data”. Van Dijck and Poell (2013, p. 10) add to the theme of data as a flow:

> Much of social media data's value lies in their real-time “live” appearance: platforms claim they can track instantaneous movements of individual user behaviour, aggregate these data, analyse them, and subsequently translate the results into valuable information about individuals, groups, or society at large.

Van Dijck and Poell (2013) are describing the code-driven editorial control over user generated content that sorts multiple data streams. This process promotes the discoverability of selected people and their content over others and repackages these streams of data into information that is fed back into the system. Repeats of this cycle lead to an exponential increase in streams of information that is disseminated to the userbase by way of product features. Raley (2013, p. 123) describes the process of a complex set of practices that enables advertising technology to tap into multiple data streams and generate user profiles before the fact. This ecosystem sustained by digital data trails is a result of collaboration between different parts of the advertising technology industry:

> Data speculation means amassing data so as to produce patterns, as opposed to having an idea for which one needs to collect supporting data. Raw data is the material for informational patterns still to come, its value unknown or uncertain until it is converted into the currency of information. And a robust data exchange, with so-termed data handlers and data brokers, has emerged to perform precisely this work of speculation.

Raley (2013) described the data collection philosophy of digital platforms wherein predictive profiles of all digital users are created and audiences are traded based on data that will be collected in the future. This immediate access to information and live insights can be converted into recommendations via feature engineering and the same data is used to create real-time auctions for advertisers. boyd and Crawford (2012, p. 663) describe big data as “an interplay of technology, analysis and mythology”. The essence here is that big data by itself does not present value for digital platforms, but its potential lies in the use of technology to derive consumer

stories from it. boyd and Crawford (2012) also challenged the definition of big data that only uses the size of a database as a distinguishing factor and described machine readability as a defining factor for qualifying as big data. Mayer-Schonberger and Cukier (2013) write about such usage in terms of conversion of actions, interactions, location and intent into data that is used by digital platforms such as Google to understand and predict future user behaviour. The ability of digital platforms and ad tech to convert big data into information that is acted on in real time form the spine of mass personalisation of advertising and recommendations.

Personalization of communication presents opportunities to target audience need with personalised creatives. This is very different from advertising in traditional mass media where a similar set of creatives are served at scale to an undifferentiated audience. Bolin and Andersson (2015, p. 7) explained that algorithm-based universal observation has moved the focus from socio-economic classification of media audiences to an algorithm-driven, pattern-based process of creating advertising-relevant cohorts:

> Broadcasters and advertisers often remain faithful to the well-worn, scattershot broadcasting heuristic rather than taking the risk of relying on a highly tailored, convoluted process of identifying 'relevant' patterns and then tailoring communication to those fractions of user profiles that emerge.

According to Bolin and Andersson (2015), the broadcast media industry relied on socio-economic classification and "representational statistics" due to the non-interactive nature of audience measurement. Broadcast media audiences are measured via representative samples that are built on demographic quotas; hence, buying and selling of audiences is limited to broad demographics. Demographics are a proxy for intent and consumption.

To understand the role of online platforms in driving sales, Calder (2009) recruited visitors to multiple websites to identify two different kinds of engagements: "personal engagement" and "social interactive engagement". Calder concluded that online media involves a distinct form of engagement due to the interactive nature of the medium. Calder found that this engagement has its own impact on advertising effectiveness. As the internet usage grew, so did the quantum of usage data. Bolin and Andersson (2015, p. 4) called the digital data trails participatory surveillance where "media audiences participate in the metrification of their habits". This participation is voluntary but not always well informed about how the software code is massaging the interaction. Pasquale (2015) outlined the issues of automated sorting delivered by software code and metrification of clickstream data and its use to grant or deny services to prospective consumers. Pasquale (2015, pp. 32–33) described the process as consisting of "a thousand eyes" maintaining thousands of "consumer scores". The visible part of this process are the ads that

inform the consumers about products and services; the black box is the algorithmic scoring of consumers to determine services and prices.

In order to understand the black box of algorithmic scoring and real-time ad serving, Berry (2011, p. 4) outlined invisible advertising technology and highlighted that software had permeated every aspect of life to the extent of becoming ubiquitous and invisible at the same time: “as software increasingly structures the contemporary world, curiously, it also withdraws, and becomes harder and harder for us to focus on as it is embedded, hidden, off-shored or merely forgotten about”. To navigate this increasingly opaque world of mediated interactions and sophisticated information environment, the users need to understand media economics and awareness of algorithmic ability to simultaneously enable and shape these interactions. Carmi (2020, p. 48) explained how technology guides the user through pre-defined paths that are invisible to the average user who only interacts with the user interface of a digital platform: “Software, code, algorithm, and protocol affect both non-human and humans, as their operations and executions direct the way that people can behave, understand and communicate with and through computational territories”. Carmi (2020) is drawing attention to the fact that interactions with the software code running the user interfaces is not a simple input–output exercise and that the output is carefully crafted to elicit a pre-determined behavioural response from the user of the platform.

The scholarship (Pasquale 2015, Andersson 2015, Berry 2011, Carmi 2020) provides the concepts for defining data literacy and algorithmic awareness required in the population to safely navigate the world of “a thousand eyes” scoring them for thousands of advertisers. Sadowski (2020, p. 45) contributes to the discussion about the ubiquitous nature of code and universal tracking for the datafication of users: “As smart tech becomes more advanced and more pervasive, we are blown apart into increasingly more streams of data.”

This section described the generation of digital data trails and the use of software code to manage this information. The ubiquitous nature of code makes it difficult for the consumer to understand it and it favours large institutional entities that have the ability to manage and harness the power of data via sorting algorithms. The next section explores scholarship around the explosion in data generated by the digital exhaust of internet usage and the evolution of postdemographic audiences.

2.2.1 User Data Explosion

News media regularly reports about user growth for the internet as a medium and individual digital platform. Business publications report on the most recent revenue projections from

advertiser-funded digital platforms. Advertising revenues and the technology that is used to deliver advertising to smartphone users can be used to understand the functioning of advertising technology (see section 1.3). As seen in the introduction, more than half of global advertising spends are on digital platforms and most of digital delivery is automated and driven by algorithmic interpretation and activation of audience data exhaust. Napoli (2014, p. 346) explored the role that algorithms play in determining the user experience in terms of search results and recommendations, but also the role of user choices in shaping the algorithms.

> Specific patterns of mutual influence (i.e., duality) in the intersection of users and user information regimes are already being identified. One important pattern, for instance, is a certain amount of reflexivity that is inherent in much algorithmically driven media consumption.

Napoli (2014) developed a perspective that humans and algorithms learn and influence each other and concluded that the users on most parts liked this adaptive nature of the algorithm with regards to content and advertising recommendations. The fuel for running these algorithms is the digital data traces generated using technology in both the online world and the offline world. Lupton (2020b, p. 46) did a critical analysis of different metaphors used to describe digital data or big data that gets generated by human–computer interaction and divided them into two sets: one set used liquid metaphors like “deluge”, “flood”, “ocean” and even “tsunami” and on the other set used excretion metaphors like “sweat”, “exhaust”, “trail”.

Raley (2013, p. 123) compared the business of data with the business of oil and says that “data speculation means amassing data so as to produce patterns” wherein data is collected and valued for its future use. Raley (2013) is explaining the “dataveillance” industry that has come up with digital platforms and data brokers collaborating to create linkages to create detailed profiles of internet users, and its use for “online behaviour advertising”. Raley (2013, p. 124) explored dataveillance in three dimensions: “Dataveillance in the present moment is not simply descriptive (monitoring) but also predictive (conjecture) and prescriptive (enactment)”. Out of the three dimensions, the predictive dimension of dataveillance is about amassing big data to predict future behaviour of digital audiences based on their past preferences. These predictions are crucial to understand real-time bidding infrastructure and programmatic advertising. Raley (2013) defined dataveillance as the layer of connecting data entities that aggregated signals from digital exhaust to identify and track individuals over time and space.

Bodle (2016, p. 139) describes the use of behavioural targeting as a way of activating advertising surveillance: “The tracking of personally identifiable data, both behavioural and transactional, depends on the ability to fix one’s identity over time so that seemingly innocuous disclosures can

culminate into a useful and valuable profile". According to Bodle (2016), advertising surveillance is as one of services resulting from dataveillance and can incorporate behavioural and transactional data from multiple entities to create ad personalization that can be delivered through digital platforms. Raley (2013) defined the mechanics of dataveillance and Bodle (2016) described the activation of the dataveillance infrastructure described by Raley (2013) to create revenue for digital platforms through mediating access. Ad personalization using personal data of users, highlights the power asymmetry between advertisers and digital platforms that have complete visibility of their consumer, whereas there is no reciprocal visibility. To highlight this issue, Bodle (2016, p. 145) outlined a distinction between panoptic and non-panoptic definitions of internet surveillance:

> Non-panoptic definitions of surveillance identify a broader notion of surveillance as a neutral and technical process. Panoptic definitions that focus on domination recognize the power that institutions, such as corporations and governments, have over resources that put individuals at a massive disadvantage.

Bodle outlined a framework to examine internet surveillance through power asymmetry between users of the internet and organizations or governments, as well as the non-human nature of internet surveillance that results in customised advertising. Bodle (2016) contrasts the benefits of customised information that can be delivered by through advertising by Facebook, but is critical of algorithms acting as gatekeepers to information and filtering contrarian views in order to maximise engagement and time spent on the platform.

Re-examining the history of data collection, there is scholarship that has documented this process of the collection of transactional data into big datasets to be sorted for the purpose of profiling. Since the early days of computerised record keeping, the finance and banking industry had created databases that sorted and profiled consumers based on credit card transactions, credit histories and bank accounts. Most of the retail loyalty programs were offline, but they were used to sort and profile a set of consumers to generate competitive advantage for corporates. Gandy (1993, p. 15) in his book *The Panoptic Sort* described the use of "complex technology" to perform three interrelated processes: "identification, classification and assessment" of information about individuals that is generated as they live their routine life. While Gandy was describing the availability of existing computerised databases across different industries and the computational capabilities of organizations to profile and sort the consumers, the increased adoption of the internet was creating a similar computerised database of user behaviour in the online world. The growth of mobile internet usage after the launch of the iPhone in 2007 accelerated the capabilities of digital platforms to create databases of online behaviours like the one described by

Gandy (1996). Sadowski (2020, p. 101) draws a line through time to connect contemporary practices of advertising technology to the deployment of a panoptic sort: “the data harvesters of today are part of this lineage, but with supercharged abilities to amass, analyse, and apply data.”

Beer (2009) imagined the upcoming Web 2.0 and that it would provide a platform for users to collaborate and engage in online social interactions. Beer (2009) also imagined that Web 2.0 would evolve from being a mediator to be embedded in the life of users and there would be rise in algorithmic decision making leading to the sorting of individuals for the purposes of messaging and access to products and services. Gandy’s offline panoptic sort was in the process of transitioning into the online world through Web 2.0. Gandy (1993) and Beer (2009) were able to outline the process of datafication of human activities and its use for dataveillance. The difference between dataveillance in the pre-internet era that Gandy (1993) wrote about and the post-internet era that Beer (2009) is describing is that in the pre-internet period datafication did not offer any benefit to the people whose data was being collected and analysed. In the post internet era digital platforms offer a desirable benefit prior to collecting the data and most of dataveillance is geared towards providing product features that enhance user experience resulting in increased engagement.

Another deployment of algorithmic sort described by Beer (2009) is for advertising surveillance resulting in personalised advertising. Advertising took a leap during the formative years of Web 2.0 and moved from traditional demographics into postdemographics as digital platforms created capabilities to integrate linkages between online behavioural data with offline transaction data. Rogers (2009, p. 1) defined postdemographics as “the study of the data in social networking platforms and in particular how profiling is or may be performed”. Essentially, postdemographics was the promise of the advertising technology to provide a rich insight-driven audience profiling and targeting system for advertisers. Digital platforms delivered this promise and could provide a very detailed description of audiences and their affinities. Lee (2011, p. 434) provides the structural difference between television that is dependent on traditional advertising, planning and buying ecosystem and Google, where everything is integrated into its advertising platform: “One major difference between network television and Google is that Google vertically integrates the search engine, the advertising agency, and the ratings system”. Grimshaw (2017) was heralding the age of personalised advertising and announced the decline of audience segmentation and targeting purely on the bases of age and gender (broad demographics).

There is extensive literature about the personalisation of digital services and use of digital data trails; however, there is a need for advertising research studies to catch up to rapidly changing practices in advertising. Research studies to understand the impact of personalised advertising on

the users of digital media are scant across the globe and rare in India. Bodle (2016, p. 138) outlines the reasons for the lack of academic research studies in advertising:

> Online advertising practices have outpaced critical advertising studies due to the changes brought by an online environment characterized by ubiquitous surveillance (tracking, monitoring), big data (mining, collecting, profiling, sorting), participatory cultures (user-generated content production and social sharing), and behavioral advertising (serving, personalizing, targeting).

The very reasons that hinder academic research into the ever-changing landscape of advertising technology contribute to the successful creation of postdemographic audiences. Bodle (2016) theorises about the surveillance aspects of advertising and expects informed users of Facebook to be opposed to personalization and brings forth "the power asymmetries" between a big corporations like Facebook and individual users. Apart from the reasons enumerated by Bodle (2016) for a lack of research in advertising, there is a felt need for studies in CDS focussed on everyday data experiences. Kennedy (2018) argued that most of the data activism that is trying to shape debates around the use of consumer data is driven by "elite technical actors" and made a case for including the experiences of ordinary people with regards to datafication. If ordinary people-centric data studies are not undertaken in sufficient numbers globally, they are extremely rare in India.

Turow (2011, p. 90) wrote about the early adoption of customization by advertisers enabled by availability of data brokers who could track audiences across the internet.

> The social and consumer discrimination that defines personalized advertising results from three converging developments: advertising practitioners' infatuation with data about online audiences, the rise of companies that can provide that data in a readily accessible form, and the growth of technologies that can selectively serve advertising to individuals based on the data associated with them.

Turow is describing the building blocks of advertising technology deployed by advertisers and its revenue-generating capabilities for digital platforms. This revenue generation subsidizes the cost of access for their audiences and enables an industry of data brokers who could track audiences across the internet and create postdemographic audiences. Kaplan (2020, pp. 243–244) argued that:

> Facebook, at least, does not sell data but rather offers advertisers free access to a highly refined data selection interface as a means of targeting ads, which are in turn sold via

> competitive auction that pits advertisers against each other within a marketplace monopolized by Facebook itself.

Kaplan is describing the real-time bidding wars that happen over every available impression on Facebook where the attention of a postdemographic audience is traded. Kaplan is also bringing up an important point that unlike television the real-time bidding marketplace for Facebook is not managed by a third party.

This section outlined the history of datafication of individuals and the predictive nature of algorithms that aim to drive a pre-determined response from these individuals. The review moved to creation of postdemographic audiences from exploding digital data trails and their use to generate profitable business outcomes for digital platforms. The next section explores scholarship about the commodification of this audience. The next section reviews literature around the mechanics of the trade of audiences as a commodity between digital platforms and advertisers.

2.2.2 Audiences as Commodities

As seen in the previous section, there has been a rapid rise in the adoption of advertiser-funded digital platforms (AFDP) and delivery of programmatic advertising powered by predictive analytics running on a vast amount of digital data trails. In the digital world, a user generates a huge amount of data related to moments, locations, interests and fitness. Young (2017, p. 207) compares broad demographic data like age, gender and location that was available for traditional marketers to rich contextual data from digital platforms that can predict affinities for travel, cities, hotel and airlines to name a few. Young (2017, p. 207) described the consumer as “leaving an observable, trackable exhaust trail of behaviour and attitudes, coupled with the technology to observe, collect, and curate these brand experiences bespoke to the individual”. This data gets collected and stored in an indexed computerised retrieval system to create the audience commodity that is traded between advertisers and digital platforms.

The concept of audience commodity was introduced by Smythe (1981) while initiating the blind spot debate. Smythe (1981, p. 233) called this commodity “audience power” that is created when media audience watch commercials and in some instances are influenced to make purchases of the advertised product or services. This debate was taken forward by Meehan (1984) who contended that the commodity that was traded was the estimation of a number of people watching the television commercials. Meehan (1984) brought in the measurement perspective to the blind spot debate and described the audience ratings produced by Nielsen in the USA as the commodity that was traded.

Audience power as described by Smythe (1981) is quantified by ratings (Meehan 1984) and these ratings determine the popularity of the audience and are also used as the currency that is used by advertisers to buy time-weighted reach or ratings, with the aim to reach out to a desired number of audiences. Caraway (2011) disagreed with Smythe on the basis that the buyer–seller relation is between platforms and advertisers and the audiences that are being sold are not a party to the agreement or any commercial terms.

> In truth, the audience does not appear as a seller of a commodity in Smythe's formulation. The actual transaction is between the media owner and the advertiser. Even though the audience commodity was constructed to resemble labor power, the owner of the commodity is not party to the transaction.

Caraway's view is an extension of Meehan's argument that the trade happens between the media platform and advertiser. The viewer does not seem to have a direct role in the trade. Based on these arguments, it can be concluded that advertiser-funded media platforms have two sets of customers, the viewers who have demographic characteristics and spend time consuming the content offering of the media platform, and the advertisers who desire to have a conversation with viewers of a certain demographic. The advertisers aim to achieve their business objectives through mediated access with the viewers and ratings are the currency that is used to buy and sell the attention of viewers. Meehan and Torre (2011, p. 70) took the debate further and argued for the need of a currency to enable trading between television networks who aggregate the audiences and the advertisers who want to communicate with their prospective consumers.

> Advertisers' inflexible demands for consumers drive the market in which ad agencies compete for exclusive contracts and the market in which audience assemblers compete for advertising dollars. For agencies and networks, the problem is how to demonstrate their effectiveness in reaching people with the disposable income, access to retail outlets, and desire to buy name brands whose prices are inflated by branding and advertising. For advertisers, the problem is how to evaluate agencies' and networks' necessarily biased accounts of their success.

Meehan and Torre (2011) define the market in terms of legal basis, economic relationships, and oversight from regulators and write about the audiences as the product of the television industry in the US. They also describe the crucial role for organizations like Nielsen that had invested in building audience measurement infrastructure that produces the currency for trading the audiences. Television audience measurement companies act as a neutral third party that provide estimates of the size of demographic audiences and their viewing behaviour through ratings. Advertisers buy the most relevant audiences depending on the product categories in which their

brands operate. Using their media buying power, advertisers exert some control over the programming that is made for the television networks as national networks try to attract audiences that are most desirable for the largest advertisers. The trading relationship between digital media platforms and advertisers depends on similar factors; however, the role of third-party measurement agencies gets diffused due to the global nature of the userbase of digital platforms.

Fuchs (2012, p. 708) concludes that there are three elements in the exploitation of digital labour: 1) Coercion, 2) alienation and 3) appropriation. Fuchs' conclusion imagines the online world as monopolies with no competition whereas there are advertising-free options available to almost all advertiser-funded digital platforms. For example, YouTube, the largest digital video format, has an ad free version of YouTube premium that has over 20 million subscribers. Alienation describes corporate ownership of the digital platforms; however, these global digital platforms require investments in employees, offices, marketing and data centres to name a few, and these investments are made before the users commit any time, advertisers commit any money and before any revenues come into the coffers.

The key argument for appropriation is that users do not have ownership or control over their data; however, all major social media sites offer an option for users to download and delete their data or their profile and exit the platform. All major digital platforms publish user data retention policies and all the data stored by organizations like Google and Facebook is periodically deleted and the periodicity can be set by the user. Fuchs (2012, p. 712) develops a dystopian view of digital advertising and writes that "most online time is advertising time", whereas most of the successful digital media platforms offer some level of control over advertising experience. Fuchs (2012, p. 712) argues that most of the advertisements are "about products of powerful companies", whereas the majority of advertisers on digital media companies are small and medium businesses that cannot afford to advertise on television.

As more and more interactions are mediated via digital platforms, it becomes increasingly difficult for everyday users to detect the presence of code let alone understand its impact on interactions. These interactions create the currency described as "audience commodity" by researchers in this section and this debate has evolved from covering the commodity created by broadcast media to the contemporary digital audience commodity. The multiple dimensions of the commodity that is traded between digital platforms and advertisers are: the ability of media to influence buying behaviour of its audience, also called audience power, of the size of the audience that can be reached by an advertising campaign or ratings, and the ability to create and target a postdemographic audience that combines traditional demographic information of age and gender

with interest, intent and location. To successfully negotiate the complex interconnected data streams and maintain privacy, it is important for the audience to understand and develop data literacy and there have been multiple studies to define and critically analyse data literacy among the users of digital platforms.

2.2.3 Data Literacy

This section investigates scholarship around data literacy and the level of data awareness and privacy consciousness among users of digital platforms with a focus on accessing the internet via a mobile device.

Digital literacy predates data literacy and According to Bawden (2008), the term digital literacy was first introduced by Paul Gilster in 1997 but did not gain prominence until about a decade later. Bawden (2008) described the initial approach to understanding digital literacy as the interaction of people within the new digital information environment and whether digital literacy was a technological issue or creating a framework to navigate digital communications and interactions.

Since most of subscription free digital platforms are advertiser-funded, it becomes imperative that all the data generated by users is warehoused in relational databases and exploited to provide the necessary signals to advertisers enabling them to target the most suitable audiences. Whitaker (2000) describes large interconnected relational databases as “inferential and predictive” and their use in creating personalised advertising for advertisers. Apart from advertising, large relational databases are also used for welfare surveillance. Gilliom (2001) traced the history of the use of data by the welfare state to sort the citizens receiving state benefits and ever-increasing requirement for information to be included in the positive list. Gilliom highlighted the importance of understanding the implications of individual data and how it impacts all aspects of people’s lives. Andrejevic (2014) used the digital divide to highlight the power imbalance between the data controllers like governments and corporates that had the resources to gather and analyse data about users of digital platforms and the users that generate such data but neither have visibility or control over their own data. Current literature has moved on from digital literacy to data literacy, as all the major digital platforms are building their businesses and profits through delivering personalised advertising using digital data trails generated by smartphone usage.

Carmi (2020, pp. 191–192) writes about the distinction between control over consumers’ visibility in consumer to consumer (C2C) settings and contrasts it with consumer to business (C2B) settings

where the consumers have little or no control over their visibility. An example is the visibility of individual cohorts to advertisers but a lack of visibility of the advertisers to the consumers.

> So, while they are offered more control over which other individuals within their network see their content, the control they have over the back-end aspects of the interface is quite limited. People cannot control if Facebook and other third-party companies listen to their behaviours, because they are not offered such an option.

Data literacy indicates consumer awareness of their visibility in C2C and C2B settings and the required expertise to exercise control in both the situations. According to Carmi (2020), the issue for user data literacy is that the underlying technology is largely invisible and most of the users are not willing or able to invest the time required to gain the required expertise to fully understand the use of their data beyond the most obvious. Pybus et al (2015) used a mobile application developed for mimicking data gathering practices of mobile operating systems as well as digital platforms to develop pedagogies to deliver tools about data literacy to the general population. Pybus et al (2015) discovered a good level of awareness about the data gathering practices of digital platforms. The application of this knowledge to exercise control was hampered by the fear of missing out (FOMO) on friends in digital spaces and prevented them from migrating to other platforms. The main contribution of Pybus et al (2015) was in creating a collaborative framework for creating datasets that could be accessed by researchers and users alike. In most of the cases, the users are not aware about the type and kind of information that is being gathered and even if it is mentioned in the detailed terms and conditions, the users find it difficult to understand the various dimensions of the consent that they are providing.

Another study highlighting the prevalence of data literacy in Australia was conducted by Nguyen and Solomon (2018) and it explored issues like data awareness and its implications for everyday users of digital platforms. This pan-Australian study by Nguyen and Solomon (2018, p. 3) found that users of digital platforms "did not fully understand specifically what types of information were being collected and shared about them; Experience barriers to reading Privacy Policies; and/or lacked genuine consent over the terms upon which they provide their information". Through this study Nguyen and Solomon (2018) also identified seven modes of passive data collection: 1) cookies, 2) web beacons/pixel tags, 3) device information and tracking, 4) fingerprinting, 5) facial recognition, 6) payment cards and loyalty cards and 7) data collection from third parties.

Pangrazio and Selwyn (2019, p. 428) propose a framework for personal data literacies with five domains: "(1) Data Identification, (2) Data Understandings, (3) Data Reflexivity, (4) Data Uses, and (5) Data Tactics" to research the subject. Pangrazio and Selwyn (2019) suggested that this

framework can be used for developing tools for the individuals to better understand their data environment. The framework presents an increasing level of understanding about personal data that can be possessed by an individual and can be used to develop and evaluate data literacy programs. An individual can be evaluated to understand their current state with regards to the five domains and target resources to help them move to the next domain. The first three domains build on the understanding from the previous domain, whereas the last two domains represents an evolved understanding of data circumstances and the willingness and ability of the individual to activate this knowledge into action.

The personal data literacy framework provides a toolkit to undertake a critical analysis of various degrees of awareness, starting from a basic understanding of the data that users provide to the platform to a higher-level awareness where the users try various tools to manage and control their data privacy. These two levels of data awareness (domain 1 to 3) and activation (domains 4 and 5) can be used to characterise users into data aware and privacy active.

Pangrazio and Selwyn (2019) found that youth (aged 13 to 17) were concerned with the interpersonal aspect of their data on social platforms like Facebook and potential misuse of their data by unknown actors. However, even after the respondents were exposed to the amount of data that a digital platform can collect about them over the course of a year, they still did not feel compelled to change their behaviour as being on social platforms was something that they could not do without and felt powerless to make the change. Pangrazio and Selwyn (2019) found that "most saw the sharing of data with advertisers as an acceptable trade-off". These "digital natives" could not come up with a viable strategy to counter the data gathering practices and were consciously aware and accepting of the data and information environment. The participants of this study had data literacy confirming to the first two domains of data awareness and understanding but had not graduated to the last three domains.

Post 2019 there have been sustained efforts to study users and their data practices and University of Liverpool has been working on the Me and My Big Data project. As part of this project, Carmi et al (2020, p. 10) have developed a theoretical framework of "data citizenship" consisting of three areas, "data thinking", "data doing" and "data participation" to critically understand the strategies adopted by an individual to navigate the data rich digital platforms encountered in all walks of life. Carmi et al (2020) also talk about the differing level of understanding in how media in general and digital media in particular are funded and the differing literacy across various socio-economic statuses in society. An understanding of data collection and usage is linked to a better understanding of the business models of the digital platforms that users consume on a day-to-day basis. Carmi et al (2020) introduced the societal aspect of data engagement by introducing

"network of literacies" and provided the framework to interpret individual data practices based on their social cohort.

This section highlighted the lack of research into the impact of advertising technology on data experiences of everyday users. Studies reviewed in this section outline user perspectives on data exchange with digital platforms and the trade-off of convenience and privacy. The next section reviews scholarship related to digital data privacy and its implications for everyday users.

2.3 Privacy

There has been increased coverage of issues surrounding the privacy of individuals in the digital world, in the news media. The issue of data privacy predates the global adoption of internet technologies. Gilliom (2001, p. 12) wrote about "Carnivore", the internet surveillance system uncovered in 2000. Recently there are examples of using consumer data privacy concerns by Indian and US governments to potentially ban mobile apps like TikTok that are owned by Chinese corporations (see section 1.5).

In the case of AFDP, users trade their usage data as well as pay attention to advertisements in exchange for free or subsidised access to desired services. Van Dijck (2013, p. 170) outlined that "a user is simultaneously a resource provider, a product, and a customer. Buying into the 'free' deal, users barter away privacy for convenience and facilities." There is an acceptable benefit for the user of the software service and in most instances the privacy aspect of the exchange is obfuscated in lengthy terms and conditions and the user is largely apathetic to such issues. Sadowski (2020, p. 56) describes the contracts that consumers sign to indicate their consent via an end user license agreement:

> EULAs are known as "boilerplate" contracts because they are generically applied to all users. They are one-sided, non-negotiated, and non-negotiable. They are long, dense legal documents, designed not to be read. You either unquestioningly agree or you are denied access.

The role of these lengthy End User License Agreements (EULA) is to ensure compliance. The end goal of a digital platform is to ensure consent for the exchange from all its users. These consents also include possible usage of data and data joining with owned and non-owned sources of information. The role of advertising technology as defined in this exchange is to ensure that the data that is collected as part of this exchange is used in creating the audience commodity that is desirable to advertisers. According to Einstein (2017, p. 159), when audiences are using a free digital service, there are "multiple advertisers bidding at auction for attention". There are layers

of complexity in the trade for audience attention. West (2019, p. 31) described the privacy compromise undertaken while using e-commerce services like Amazon and its voice assistant Alexa: “Fundamentally, Amazon offers to serve us by knowing us, including the domestic, private side of ourselves represented by our product searches, our purchases, the media we consume, and now with Alexa, what we say and how we say it.” The premise used by Amazon is that information sought from users is legitimate in nature and is used to perform the desired service of enabling an e-commerce transaction on the platform or listening to conversations to be of assistance.

The business success of AFDP lies in deriving value from the commodity that they are selling. The commodity that is being traded is the attention of qualified audiences with the complete process occurring in real time. The qualification of audiences for every advertising opportunity is enabled with profiling data that is passively captured in the online world and combined with offline databases. The real-time bidding of audiences requires a pre-existing digital dossier where audiences are created before the fact and advertising opportunities are predicted in advance to set up a real-time auction. The process of creating individual digital dossiers is discussed in the next section.

This section shows the growth in generation, storage and analysis of large-scale consumer databases and their use for advertising surveillance and personalised advertising. This creates issues of data privacy for users, who have to engage in privacy compromises to be able to access the services offered by AFDP. The next section explores the collation of individual data trails into digital dossiers leading to mass customization of advertising and its impact on data privacy.

2.3.1 User Privacy Management and Concerns

Data collection about the consumers has been practiced well before digital platforms gained prominence. The ability of the corporates to willingly co-opt their customers into sharing their data was dependent on the value that customers assigned to the benefits that were accrued by them in exchange. Davies (1997) credits the overall level of anxiety in society about crime and unemployment leading to the rise of “public interest” where citizens volunteer to share their personal data with government agencies. Similarly, consumers also willingly share their details with corporates when the requirement is expressed in an affirmative manner. According to Davies (1997), an example of such neutral language that mitigates consumer resistance is when “customer profiling systems” by supermarket chains were promoted as “loyalty clubs”. The main contribution of Davies is to document the data practices of retailers in the past decades. These databases represent the offline transactions of consumers before the proliferation of digital

platforms made it possible to collect behavioural data at a national and global scale. Davies (1997) examined the participation of consumers in the creation of offline relational databases.

Westin (2000, p. 8) defined privacy as "the claim of an individual to determine what information about himself or herself should be known to others". This lock and key or a walled garden approach to privacy is also described as a "secrecy paradigm" by Solove (2004, p. 42) who wrote that "privacy is about concealment, and it is invaded by watching and by public disclosure of confidential information". People are willing to let go of this shroud of secrecy around their data if presented with a reciprocal benefit. This benefit is the perceived sense of enhanced safety as described by Davies (1997) or some preferential treatment by corporates. Phillips (2004, p. 693) viewed privacy through the prism of four concerns: "'freedom from intrusion', 'negotiating the public/private divide', 'identity management' and 'surveillance'". Phillips (2004) delves into privacy enhancing technology as a tool for mitigating privacy concerns leading into discussions about regulation and definitions of identity.

James (2009) studied the participation of American youth on various social platforms and discovered a broad spectrum of behaviours, ranging from active management of user privacy to a complete disregard of privacy protections offered by social platforms. James (2009, p. 40) writes about possible peril when the youth fail to understand "the persistence of, searchability of, replicability of, and invisible audiences for the information that they share about themselves online". James is suggesting that there is a certain section of youth who do not use the privacy settings provided by social networking sites to limit the visibility of their posts and could expose themselves to the permanent storage of their personal information that is searchable by anyone without a time limit and reposted any number of times by unknown individuals or actors.

Turow et al (2015) investigated the combination of offline databases created through loyalty programs and ways in which this data is combined with data from third party data brokers to "reshaping the shopper", "reshaping the store" and "reshaping the deal". Turow et al (2015) described the enrichment of offline purchase data coming from loyalty schemes, with hundreds of additional data points from information brokers like Acxiom, Experian and via technology like Bluetooth Low Energy (BLE) beacons. Additionally, big box retailers and supermarkets use passive data, gathered by financial giants like Mastercard and Citigroup and location providers like FourSquare, to reach out to customers beyond their stores with customised deals, offers and experiences. This is another example of the deployment of neutral language combined with a reciprocal benefit that encourages users to trade their data.

The issue that emerges from these discussions is that privacy gets compounded into a network of relationships between multiple entities rather than remaining as a one-on-one relation between

an individual and a corporate. Mayer and Cukier (2013, p. 176) provided an alternative framework of privacy in the "era of big data" where the value of data is not at the time of collection, but in its secondary usage. They propose to move from "privacy by consent" to "privacy through accountability" where the regulators hold the data use accountable for the secondary usage of consumer data. A one-on-one relationship is easier for an individual to understand as compared to a decoding the network of relationships that form the backbone of ad tech. Mayer and Cukier's (2013) framework shifts the responsibility of guarding the privacy of data on to the data aggregator rather than the user. However, privacy is often not the key deciding factor when the consumers choose between competing technologies, retail stores, goods and services or digital platforms.

2.3.2 Digital Dossiers and Personalisation

As discussed in the previous sections, smartphone usage generates a digital data trail and it is stored and analysed to create a variety of information that is critical to creating customised advertising. Data from the online and offline world is joined using common identifiers deployed by advertising technology. These processes for aggregating user data and running algorithms to build qualified audience for achieving business objectives were developed alongside the gain made in computing. Gandy (1993, p. 17) wrote about "assessment" as the next step after the audiences and their data has been classified and involves the examination of probabilities of profitable outcome in response to an intervention. In the case of advertising technologies these predictions are about future advertising opportunities and assessment is before the fact profiling of audiences into targetable cohorts.

The architectural ethos of ad tech systems is similar to Gandy's description of panoptic sorting. Advertising technology aggregates information from disparate databases to identify and track individuals based on their digital trail data and classifies them into audiences. In the early days of mobile internet, Thrift (2005, p.220) wrote about the evolution in computing from being location-dependent to mobile and all pervasive: "This means both that devices will become more location-aware, knowing where they are in relation to users and other devices, and that they will be able to interact and communicate with and adapt to users and other devices". This location awareness of devices, combined with open nature of Android mobile operating systems, offered a window of opportunity for governments and private organizations to push the limits of individual privacy in the digital world.

Whitaker (2000, pp. 125–126) had defined the tracking aspect of advertising technology two decades ago, while detailing the shift from surveillance to dataveillance using digitised databases:

> One of the crucial elements in the transformation of data into a commodity is the practice of data-matching or data linkage, whereby separately collected and separately organised pieces of data are matched or linked to produce new and valuable information.

In the case of ad tech this process of audience identification, classification and segmentation happens through an invisible backend and the consumer is only exposed to the resultant advertising. Data about individuals is combined from multiple sources to create “data doubles” and algorithms are deployed to predict the user preferences and generate recommendations. Solove (2004, p. 2) defined the process of the creation of data doubles: “Data is digitized into binary numerical form, which enables computers to store and manipulate it with unprecedented efficiency”. Algorithms act on these databases to create information that drives the economy of digital media companies. Taken together, these four authors over a period of a decade had described the contours of advertising technology and documented the creation of data doubles. They addressed the issues of combining consumer's behavioural and transactional data without their informed consent, the creation of digital dossiers and the functioning of advertising technology. Andrejevic (2007) while discussing lateral surveillance also prescribes self-surveillance: "in many cases, unless we start monitoring our publicly available data shadows, we're often unaware of the information available about ourselves online." The main issue with ad tech was that the data shadows used for algorithmic sort were not publicly available. In the next decade the scholarship focussed on the information generated from these digital data dossiers and its use for advertising surveillance.

Cheney-Lippold (2011) wrote about the role and power of algorithms in the classification of user data gleaned from multiple sources into highly accurate information about individuals. Centrality of information for digital platforms was detailed by Mayer-Schönberger and Cukier (2013, p. 96): “with the help of big data, we will no longer regard our world as a spring of happenings that we explain as natural or social phenomenon, but as a universe comprised essentially of information”. Mayer-Schönberger and Cukier (2013) described the world where data is generated by a multitude of devices and the ability to convert this data into usable information and the flow of information will have the biggest impact on how the information ecosystem of digital businesses will progress.

In the past, Castells (2007, p. 61) had called this the “Information Technology Paradigm” and proposed that “information is the raw material” of this paradigm. The information paradigm shines a light on the role of code in aggregating the physical and virtual activities of individuals and at the same time works on this data to derive actionable intelligence. This intelligence is

activated by advertising technology to create target audiences before the fact and the process leaves the users unaware of the existence of algorithmic profiling. Mayer-Schönberger (2009, p. 137) espoused a pragmatic approach to data privacy, that data sharing in the "interconnected world" is not an option but is a question of how to manage the process: "It is a right to determine primarily how – rather than whether – one participates in sharing information". This pragmatic approach was supported by the participants of Lupton's (2021, p. 16) research: "Australian participants demonstrated pragmatic acceptance of the inevitability of practices such as data profiling and targeted advertising, and the probability that datafication and dataveillance will accelerate in the decades to come." Lupton discussed the importance of the voice of young people and their experiences with digital platforms and how their acceptance of data practices was very different from the scholarship alarmed at datafication and heightened dataveillance.

Sandel (2012, p. 12) described the moral limits of markets and argued that "certain good things are corrupted or degraded if turned into commodities" and warned against sliding from a "market economy" to a "market society" where everything is for sale. This distinction becomes important in understanding the commodification of audiences and the "privacy divide" that has emerged amongst those who can pay extra for the added layer of privacy in the digital world versus those who cannot afford the payments and are left exposed in the digital world. This linkage completes the data dossier and advertising technology makes it accessible for advertisers. This process of linking online and offline data is explained by Gillespie (2014, p. 8): "Beyond knowing the personal and the demographic details about each user, information providers conduct a great deal of research trying to understand, and then operationalize, how humans habitually seek, engage with, and digest information". Gillespie is describing how postdemographics is activated by information-rich digital media organizations to classify their users into advertisers' target audiences. Similarly, Van Dijck (2013, p. 5) wrote about digital platforms and their usage becoming a quotidian habit and the usage in turn enabling the platforms to gather large datasets that could be converted into advertising signals for advertisers.

Carah and Brodmerkel (2020, p. 6) discuss the real-time nature of advertising technology: "digital media though allow for consumers and cultural participants to produce representations of the brand that are seen at scale and for consumer innovations with the brand to be monitored and leveraged in real time". Carah and Brodmekel investigated the manner in which brands are increasingly embedded into digital interfaces and help consumers co-create the brand narrative. Worldwide reach of digital platforms has enabled brands to engage with consumers using advertising technology in a way that was not possible in the offline world, where the consumer was usually at the receiving end of one-way communication with limited opportunities to engage in two-way communication.

People going about their daily life move seamlessly between activities in the physical world and the usage of digital platforms. Advertisers are interested in their life stages and consumption patterns to develop an ability to communicate their messages to the appropriate target audiences. Advertisers, media platforms and data brokers enabled by a sophisticated layer of advertising technology that is powered by automated storage, processing and linkage of data across diverse databases are able to deliver the right message at the right purchase moment to drive their business objectives. Carmi (2020) explained how cookies form the invisible underlying communication system that drives the digital ad tech engine and provides it the ability to identify individuals and collect passive data about their online activities.

> With this special restructuring of the online territory people's experience on the web was conducted in a specific space called the "front end", while the advertising industry's activities were conducted in the "back end". This created a knowledge boundary between "average" users and the online market which was operating in accelerated rhythm at the back-end.

The consumer is only exposed to the user interface of a digital application and is generally not aware of the use of advertising technology to track them across the internet. As humans exist and transact in the physical world, there is a requirement for ad tech to link their digital data with offline data about location, movement and purchases in the offline world. The intersection where the data linkages happen is the point where the consumer loses control over the privacy of their data. Rich consumer segments and real-time targeting capabilities enable marketers to have dynamic conversations that can be automated to pre-described events and provide seamless integration into users' digital interfaces. Carah and Brodmerkel (2020, p. 9) describe these seamless entanglements between users and brands: "The digital, participatory media environment provides marketers and advertisers with new and innovative opportunities for creating strategic and commercially-driven consumer experiences". Branded communication has evolved from a passive pre-packaged advertisements to active ingredients that are immersed into consumer experiences and are customised to their personal situations. Consumer disquiet at this invisible layer to tracking technology led the research participants of Steedman et al (2020, p. 826) asking for a " strong regulatory framework for data protection".

India is primarily an Andriod market and the penetration of Apple devices is negligible. In most of the western countries Apple products outsell Android devices. This difference in predominant operating systems impacts the ability of individuals to manage control over their own data. The discussion returns to the issue of the privacy divide as discussed by Westin (2000) above. Investigated on a global scale, there is an existence of a "privacy divide" between the wealthy

Western world and developing economies like India. It is possible for people in the developed world to afford instruments like an iPhone that are at the forefront of user privacy. Apple has been championing user privacy through its efforts to prevent cross-site tracking of users on Safari for the past few years. Its Intelligent Tracking Prevention is a privacy mechanism that prevents users from being tracked across the internet through restrictions on sharing of apple advertising ID (IDFA). This issue has been widely reported in the news media but has been largely ignored in academic research.

Solove (2004, p. 82) described the privacy paradox that exists amongst consumers where they express concerns about their privacy but do not make purchase decisions based on privacy protection. “Companies only rarely compete on the basis of the amount of privacy they offer. People often do not weigh privacy policies heavily when choosing companies. For example, people rarely choose phone companies based on their privacy policies”. A long history of privacy laws in Western nations and an absence of privacy regulation in India also fuels the privacy divide. There is little or no incentive for corporations to differentiate their product offerings purely based on user data privacy. Due to global product standardization some digital platforms roll out privacy features designed to satisfy European or US regulators in developing countries. In the absence of a legal requirement in India, global technology companies tend to offer minimum protections that comply with local laws and other features that they cannot turn off by location.

Shöenberger (2017, p. 304) related such an exchange to a “privacy paradox” where there are voices in the media and surveys that talk about huge concerns around consumer data privacy, but the actual behaviour of the users is totally the opposite. Steedman et al (2020, p. 826) discovered a complex relation between the perception of an organization and its impact on trust in its data management capabilities, where the participants "agreed that a PDS is better than the current model, because it offers more control to individual users" but were circumspect on the aspect of time and security. They discovered the same privacy paradox as outlines by Solove (2004), Shklovski et al (2014) and Shöenberger (2017). Rapid adoption of mobile internet and a lack of privacy protection regulation prompted Arora (2016, p. 1694) to label India as the “Global South” of the digital and privacy divide compared to Western countries where there is growing recognition for protecting the online privacy of its citizens, calling it the “Global North”. A similar North–South divide in terms of access to data privacy was described by Westin (2000) while introducing the concept of the privacy divide between the upper class and the lower class in society:

> At the sociocultural level, the real opportunities people have to claim freedom from the observation of others are shaped by environmental factors, such as crowded cities, and

> class factors of wealth and race. In this sense, privacy is frequently determined by the individual's power and social status. The rich can withdraw from society when they wish; the lower classes cannot.

As scholarship about data aggregation and the limits of individual data privacy has evolved, so have the business practices of digital platforms. These rapid and continuous developments have been highlighted by Gurses and van Hoboken (2017) who credit the adoption of agile methodologies of information products for the ever-evolving nature of advertising technology. In other words, advertising technology infrastructure is always functional but it's never completely built.

This section examined scholarship related to the capabilities of advertising technology in tracking and connecting individuals across online and offline spheres. The concept of the privacy divide across geographies and social class was also examined. There is concern amongst scholars that individuals are trading privacy for access and in turn enabling aggregation of postdemographic profiles without explicit informed consent. This aggregation of online and offline behavioural and transaction data and general user apathy is discussed through the concept of the privacy paradox. Scant literature about these issues in an India-specific context comes up and so does the position of India being the Global South, with the least amount of public debate or policy initiatives towards data privacy. The next section brings together the debate about privacy with a review of definitions and its development over the years when users migrated from broadcast media to digital media.

The implications of data privacy concerns are manifested in the concerns about the use of data for the purpose of surveillance. The next section reviews scholarship about the ability of digital technology to surveil its users. Literature exploring the ability of users to comprehend the data management practices of digital platforms and mechanisms deployed to understand the issues surrounding it will be reviewed. The next section also explores scholarship about how the "digital exhaust" of behavioural and transactional data is combined to create audience profiles and its use for advertising surveillance. The next section examines contemporary themes in advertising surveillance leading into the discussion of participatory surveillance.

2.4 Surveillance

This section discusses the deployment of advertising technology for the mass customization of communication and the vital role that personal data plays in this process. This section reviews literature related to advertising surveillance of users for the purpose of generating a targetable audience for advertisers. Section 2.4.1 explores the functioning of advertising technology and

literature that uses the observation framework of Bentham's panopticon to study the surveillance capabilities of AFDP 2.4.2 examines literature about how free access is enabled by signing on to advertising surveillance, section 2.4.3 reviews scholarship about an entire gamut of surveillance with a focus on participatory surveillance , and section 2.4.4 reviews literature about audience profiles and datafication. Mobile internet usage has penetrated every aspect of life in India (see section 1.2) and it is more prominent in urban India and big metros. Increased use of mobile internet has led to an increasing amount of data being generated, stored and analysed by digital platforms. The analysis and management of "big data" enables the technology companies to generate detailed profiles of their customers ranging from their travel habits, modes of transportation, home, office, education, age, gender, and relationships to name a few.

In the literature, there are constructs of dynamics between the observed (audience) and the observer (technology) that highlight the existence of power asymmetry. This asymmetry has existed since the dawn of computerised data storage and retrieval systems. Gilliom (2001, p. 2), explaining the welfare surveillance in United States, writes that the observer in a surveillance system is in a position of dominance and defines surveillance as a "routine use of personal data and systematic information in the administration of institutions, agencies, and businesses". Gilliom (2001) illustrated the power asymmetry between the state that distributed welfare benefits and the receivers of welfare benefits, who have had to submit an increasing amount of personal information to be able to receive the same benefits over the years. Over the years a similar power asymmetry has been established between digital platforms and their users.

Digital platforms have certain preconditions for accessing their features and can bar users unless they provide consent for extensive data sharing. The advertiser-funded digital medium had been conceptualised in literature way before it took shape. Turow (2005, p. 119) predicted a future where varied digital services would be subsidised by advertisers based on detailed profiling of audiences:

> The circumstances are not hard to imagine. Using detailed audience surveillance, digital marketers will be able to track the media activities of their target audience in considerable detail. To ensure that they view targeted, perhaps even customized, commercials on the web or on TV the marketers may offer audience members discounts to programming, music downloads, game networks, or a panoply of other subscription or pay-per-use activities.

The digital world of today with ad-supported platforms is very close to Turow's prediction. Advertiser-funded traditional media and digital platforms have to convince advertisers of their

ability to deliver the most suitable attentive audience, in order to attract a steady flow of advertising monies.

Data collection practices are not relegated to technology companies but have been in existence for a long time. The need to develop sustainable revenue streams has provided the imperative for digital platforms to excel in the realm of “big data” and my research is focussed on understanding the awareness of users in India about what and how their usage data is collected and their reasons to take part in this exchange of data and attention for access. The next section investigates the theoretical concepts of surveillance developed from Bentham's panopticon and the scholarship contrasting it with advertising technology.

2.4.1 Bentham’s Panopticon and Advertising Technology

This section reviews literature investigating the structure of advertising technology and scholarship comparing it with the use of observation to influence behaviour as described in Bentham’s panopticon. Foucault (2008) and his discussion of “Bentham’s panopticon” is widely used as the basis for examining digital surveillance and theorising about advertising technology and its societal implications. Foucault (2008, p. 6) wrote that “the panopticon is a machine for dissociating the seen/being seen dyad: in the peripheric ring, one is totally seen without ever seeing; in the central tower, one sees everything without ever being seen”. Foucault (2008, p. 6) described panopticon as a system where the power seems to be in the hands of the observer at the same time the “mechanism” “automatizes” and “disindividualizes” power. In this conceptual mind experiment, prisoners are made conscious of the fact that they are under constant observation. Given the design of a panopticon however, the observer remained invisible, and the process of observation could not be assigned to an identifiable person. Observation was conducted by a system that the physical design of the prison made it explicit to the observed. The feeling of being under constant observation nudged the observed to follow the rules and behaviour change is achieved by a mere sense of being observed round the clock. According to the previous section, users in the digital world are continuously observed by advertising technology through their behaviour that is recorded via their data trails and is continuously profiled (see section 2.4.4).

Mayer (2009, p. 197) added a time dimension to the practice of observation and highlighted the permanence of digital traces:

> As digital memories make possible a comprehensive reconstruction of our words and deeds, even if they are long past, they create not just a spatial but a temporal version of

> Bentham's panopticon, constraining our willingness to say what we mean, and engage in our society.

The reference was about the inability to forget or remove digital traces and hence users of digital platforms would be very guarded in engaging with the wider society through digital platforms. This is an important aspect of digital data collection, as a digital echo from past actions could have a lasting impact on an individual's ability to function in society.

Foucault (2008, p. 7) described the process for achieving different objectives for different sets of individuals using the panopticon: "One finds in the programme of the Panopticon a similar concern with individualizing observation, with characterization and classification, with the analytical arrangement of space". Other authors have expanded on Foucault's concerns and have described the digital world as a modern-day panopticon. Gandy (1989, p. 63) described an information technology-enabled panopticon:

> The new surveillance can be thought of as a form of "remote sensing" where the observer is never seen. Information is processed by unknown, faceless technicians and specialists who have no direct, personal knowledge of or concern for their data subjects.

In the digital world, the software programmer writing algorithms can be thought of as the faceless specialist who does not have any personal knowledge or emotional connection with the subject.

Dandeker (1990, p. 24) described the use of a panopticon and how bodies were classified into cohorts for management and are driven towards similar objectives:

> Discipline was exerted through a continuous, uninterrupted process of supervision of the activities of the body as per arrangement that involved the partition of time, space and bodily movements. This new bodily discipline was structured through the organization of space and time in which the collectivity was broken up into manageable sections.

Using the algorithms, these data subjects can be individually identified and tagged. No two individuals are the same, but they can be grouped into cohorts using classification filters. Borges (1964, p. 60), while talking about the "Library of Babel", made a similar observation about uniqueness of individuals: "In the vast library there are no two identical books". Similarly, Bogard (1996, p. 43) described the working of a surveillance system to control the behaviour of subjects: "It records the flow of events (motions, sounds, rhythms, performances), translates activities (of bodies, persons, groups, nations, populations, whatever) into information to more conveniently control them". Foucault (2008), Dandeker (1990) and Bogard (1996) might as well be describing

the process of converting users of digital platforms into audience groups that could be targeted by advertisers using digital platforms, to nudge them towards campaign objectives. The process is similar to how Elmer (2003), Mayer and Cukier (2013), McStay (2018) and Szulc (2018) described the functioning of advertising technology for classification, profiling and targeting of audiences.

Elmer (2003) explored the concept of dataveillance and compares scholarship about automated surveillance in a panopticon with automated surveillance using relational computerised databases that are populated through various activities undertaken by people as they go through their everyday lives. These activities included credit card activities, bank transactions, CCTV surveillance in public places, and surveillance of the workplace and involve computerised databases that can communicate with each other. Just like the panopticon, Elmer (2003) puts forth the increasingly connected nature of television viewing (Tivo) and websites that reward returning visitors by filtering content based on their past interactions. Elmer also discusses television as a means of providing feedback to the users about the popular culture that is derived through the networked databases.

Writing about websites as the new domain for panoptic practices, Elmer (2003, p. 112) wrote that "Web browser and its accompanying 'cookies' have played central roles in automating, to various degrees, the collection of Web users' personal information". Elmer (2003, p. 112) devoted an entire chapter to describing how a few lines of code, i.e., a cookie, has transformed the ability of websites to track user behaviour with server-side integration and any attempt to block the cookie punishes the user by withholding "some of the most popular interactive functions of the Web". This is a direct comparison with panoptic practices using rewards and punishments. Campbell and Carlson (2002, p. 587) did a direct comparison: "the Panopticon and Internet ad servers each employ technologies of information gathering and aggregation in a methodic effort to appraise individuals and populations for various purposes of control". Their comparison was between the ability of the two methods of observation to induce behaviour change. The crowded space of internet ad servers the comparison with a single panopticon prison does not seem equitable. On the surface these comparisons seem reasonable but the plurality of technology companies and hundreds of new start-ups being launched every year does not offer any one organization a panopticon type of control.

Mathiesen (1997) wrote a critique of Michel Foucault and emphasised the missed signals of "synopticism" that were emerging in the 1970s where the rapid development in mass media technology enabled millions to observe the few, calling it "the viewer society". This is an important distinction in the discussion on surveillance as even before the internet there were technologies that enabled a two-way system of observation. Meyrowitz (2009) specifically

explained how television brought in the culture of watching people from afar into the living rooms, but the experience was unidirectional and "largely place-bound and invisible to others". Saulles and Horner (2011) named internet-enabled mobile phones as "the portable panopticon" that carries out surveillance of users on the move as against Bentham's panopticon that was anchored in a physical space. Szulc (2018) investigated the role of social networking sites, in creating extensive profiles or data doubles of its users and in constructing their identities without their explicit knowledge.

This section discussed how scholarship has evolved over time and has incorporated various elements of Bentham's panopticon to conceptualise the role played by invisible predictive algorithms in influencing behaviour. The next section provides an overview of the concepts related to the exchange of users' consumption data for free access to services on digital platforms.

2.4.2 Free Access and Surveillance

The focus area of this section is summed up by Whitaker (2000, p. 141) as "the consumer Panopticon rewards participation". Whitaker (2000) was discussing the benefits of safety and reduced crime that the participating public sought from increased surveillance of public spaces. The philosophy of a reward for participating in the advertising panopticon can also be described by modifying Whitaker as "**the advertising panopticon rewards participation**". This recounts the reason why people generally trade their privacy for free access to services offered by digital platforms. Van Dijck (2013, p. 171) offers a perspective on free access to digital platforms:

> Many online users welcome personalized ads and customized services as the ultimate convenience, whereas others repudiate them as a blatant invasion of privacy and a "locking in" to services they do not appreciate. Depending on what ideological position you take on this issue, the original proposition of "free" is either a blessing or a curse.

She was referring to the promise of Web 2.0 for free content that turned out to be an exchange rather than a one-sided flow of services. Another issue with techno-surveillance is its ubiquitous nature. Kitchin and Dodge (2011) write extensively about the presence of technology in everyday life calling it "everyware". "Everyware" is described as a concept where mundane and routine everyday tasks would be digitised with a universal availability of computing and codes embedded into everyday objects. Concomitantly, Kitchin and Dodge (2011, p. 228) also highlighted the dangers of the ever-present system, and the possibility of it becoming a surveillance tool: "One of the main implications of the development of everyware is that it opens up the possibility of widening and sharpening surveillance". The focus here is to try and understand the balance

between the use of technology to deliver “free” services to users, enabling technology to undertake routine tasks and the possibility of surveillance systems that users neither understand nor have any control over. The literature draws heavily on “Bentham’s panopticon” to explain the potential use of digital trace data for surveillance of users for advertisers.

Another reward that the users of digital platforms receive is access to the latest product features that are offered on a regular basis. These rapid deployments of product features by digital platforms are providing software as a service. Gurses and van Hoboken (2017) investigated the rapid growth of digital platforms and successful retention of users by focussing on the changing practice of product development. Digital platforms moved away from traditional product management practices like waterfall management and adopted agile methodologies, resulting in a rapid deployment of modular software features. Gurses and van Hoboken (2017) contended that a rapid release of incremental features to a software service, compared to the time-consuming process of launching fully functional software, enabled long-term loyalty from users who got used to a continually improving product experience. These rapid service enhancements bring out challenges for privacy as users are unable to keep pace with changing privacy requirements for accessing new product features. This is made possible as software is provided free of cost as a service and ties in the users in long-term usage through rapid deployment of features, keeping the user engaged. Apart from personalised services offered by digital platforms, users also welcome topical and personalised advertising.

The process of developing consumer profiles and using them to target users towards the achievement of the business goals of advertisers is compared with a panopticon-type surveillance system designed to alter behaviour. There is also a concern about the “datafication” of users and the internet of things becoming another surveillance system. In terms of architecture, the complex relationship between users, advertisers and digital platforms is perhaps closer to the “Library of Babel” put forth by Borges (1964, p. 60). Whitaker (2000, p. 48) wrote about Borges (1964) and the vision of a “Library of Babel” that “envisages the universe as a repository of information; life is the activity of retrieving and interpreting information”. Whitaker (2000) compared the physical structure of a “Library of Babel” to the upcoming “information revolution” where all information is stored in a database with infinite boundaries. This vision seems to be realised today with the storage and retrieval of “digital traces” and their use in developing product features and delivering advertising.

In this section datafication, surveillance and exchanges are viewed in relation to each other. Scholarship related to the adoption of an agile methodology of product management and long-term service contracts by digital platforms to engage users is reviewed to explore the role of new

product features in the exchange undertaken by users. The next section contrasts the impact of rapidly increasing penetration of the internet with scholarship about voluntary participation. It will also explore the various modes of participation in digital surveillance.

2.4.3 Participatory Surveillance

This section explores the participatory nature of users' interaction with mobile apps, where consumers willingly enter into agreements with digital platforms to use their services in exchange for their usage data. Whitaker (2000) introduced the concept of voluntary surveillance through "participatory panopticon". Whitaker (2000, p. 138) had a view of the panopticon that is closer to the reality of the digital world today: "It is a consumer Panopticon based on positive benefits where the worst sanction is exclusion". Instead of punishment in a panopticon prison, the worst that can happen to a person in the digital world is that they will not be able to participate in a social experience. The reasons for participatory surveillance or self-surveillance can be understood through what Gilliom (2001) explains as the burden of surveillance that an average individual suffers even before using any form of digital platforms, saying "most people live with surveillance as a regular part of life".

Participating in a phenomenon that everyone around is involved in has become essential for social existence. This research endeavours to understand the level of comfort that users have about sharing their life on social platforms and their ability to limit the distribution of their personal information. The concept of a panopticon and surveillance from the top had been reversed with the use of information technology. At times, surveillance structures like CCTV cameras in public spaces are used to broadcast images to the public who are expected to surveil their fellow humans. Ganascia (2010) writes about "sousveillance" with regards to state versus public, where the actions of people representing the powers of state like politicians or police are continuously surveilled by the general public and their actions broadcast broadly. Ganascia (2010) also explores the concept of a "catopticon" where all the information in a surveillance system is shared universally and anyone can observe everyone else; in this case the surveillance is happening from the ground up. Over the years academic literature has focussed on the two-way surveillance system that has been enabled due to the increased use of technology.

Participation in social networks also exposes individuals to surveillance from digital platforms but it also enables surveillance from peers in the network. Humphreys (2011, p. 577) defines a "voluntary panopticon" as "voluntary submission to corporate surveillance". In her research, Humphreys (2011, p. 584) discovered that some of his respondents were willing to have an

automated GPS-linked system that stored their location, instead of manually checking into dodgeball everywhere they went.

> Both Irwin and Taylor thought that automating the Dodgeball check-in process would make it easier for users. This is part of the reason why a voluntary panopticon is so powerful. People willingly submit to surveillance for the sake of convenience.

Most of Humphrey's respondents were very comfortable with a seemingly personal nature of their location data being shared with the digital platform and wanted more technology-driven experience. Whitson (2013, p. 172) describes how "participatory surveillance is rooted in the act of sharing yourself or your constructed identity with others". Voluntary participation in digital platforms also provides health benefits where users can use them to capture data about eating, to health-related activities to outright medical measures like blood oxygen levels. Commenting on the ever-growing practice of keeping records of the self or "self-surveillance", Bossewitch and Sinnreich (2013, p. 226) wrote:

> Perhaps no segment of society has more thoroughly internalized these new dynamics than youth culture. Today, most young people (and an increasing number of adults) throughout networked society volunteer an ever-growing volume of personal data, from the mundane to the profound, using services such as Facebook, Foursquare and Twitter.

Bossewitch and Sinnreich (2013) have developed a new lexicon for understanding the information exchange between digital platforms and users and reversed the concept of top-down observation of a panopticon:

> By participating in the surveillance process (both as surveillant and object of surveillance) actively and transparently, individuals can both mediate and understand the personal information they are transacting, and mitigate the inquiry of information flow by surveilling the institutions in return.

Bossewitch and Sinnreich also called this "sousveillance" and this definition does provide a means for understanding the complex interdependence of digital platforms and its users where the flow of information is not unidirectional. Digital platforms share a great deal of information via the privacy policy and terms and conditions of data usage and these can be used to gain a good understanding of the organizations collecting personal data of its users and how the data is being stored, integrated and shared with other parties. People are on continuous display on social networking sites or other digital platforms and Meyrowitz (2009) explains how television viewing has primed people for being on display for others to watch, and that the "public is familiar and comfortable with surveillance activity after participating in a form of it themselves".

Fotopoulou (2018, p. 2) gives a definition for the people who indulge in the practice of "self-surveillance":

> The quantified self is a cultural trend whose pioneers are people from the Quantified Self community (with capital QS), who undertake a range of practices of self-monitoring, data collection, management and analysis, with the use of wearable sensors and mobile technologies, in order to produce knowledge about the self.

This section described the changing nature of surveillance where people are volunteering to be subjected to be surveilled by digital platforms and have an affirmative view of the services offered by them. This section also described the practice of creating a Quantified Self, where technology is used for tracking progress and to get customised recommendations by sharing personal data. There remains a lack of scholarship in understanding everyday experiences with regards to the processes of advertising technology and its role in shaping up users' perceptions. The next section discusses literature related to the process of media planning and the role of datafication in creating customised audiences for advertising campaigns.

2.4.4 Audience Profile and Datafication

In the case of AFDP, the user experience is designed to increase attention and engagement of the consumer and the profitability of the platforms depends on their ability to sell consumer's attention to the advertisers. Better platform experience drives usage and thus generates more consumer data, creating a virtuous cycle for the digital platforms.

Sissors and Goodrich (1996, pp. 159–161) described a media plan as a strategy document that "will include recommendations for media selection and usage, a rationale for all decisions, a flowchart to show how the money will be spent". They described media planning as the strategy to select the right media to influence a qualified audience cohort in a specified time period. Sissors and Goodrich (1996) described the two most important questions that need to be answered in a media plan as 1) What is the number of individuals in the target audience that can be reached during the campaign (reach) and 2) the number of exposures to the message (frequency) in a given time period. Rossiter and Danaher (1998, p. 1) discussed the "reach" and "frequency" of advertising campaigns aimed at the target audience as the advanced task of "media strategy". Besides knowing about the number of audience members reached, understanding the number of exposures or the frequency is equally important to understanding impact.

The next important step in the media planning process is efficiency in the selection of media platforms to deliver the message. Einstein (2017, p. 149) defines media selection as “taking the right message produced by the creatives and putting it in the right place at the right time in front of the right people”. Stokes (2019, p. 129) defines audiences as “people who are exposed to, or who respond to, media culture”. Jowett and O’Donnell (2018, p. 168) highlighted the importance of narrowly defining the audiences for advertisers: “As costs have increased, advertisers are obviously eager to maximize their potential for reaching the ‘right’ audiences”. The ability of advertisers to target the most appropriate audiences is important for the return of investment on advertising spends.

Profiling audiences has been the key purpose of audience research for broadcast media and digital media alike. Profiling audiences of broadcast media was done through sample-based (see section1.4, Evolution of Media Audience Measurement) panels and was mainly based on age and gender of the audience. Young (2014, p. 207) compared age, gender and location data that was available for traditional marketers to data from digital platforms that can have information about travel, cities, hotel and airline preferences to name a few. Young describes users today as “leaving an observable, trackable exhaust trail of behaviour and attitudes, coupled with the technology to observe, collect, and curate these brand experiences bespoke to the individual”. As this data gets collected and stored in retrieval systems the audiences become a commodity that is traded between advertisers and digital platforms. McStay (2009, p. 135), writing about audiences of digital platforms, emphasised the role of audience profiles created before the fact that enable advertising engines to reach every request to serve an ad while conducting an auction to pick the winning bid.

Computerised datasets had made profiling of individuals on the basis of behavioural and transactional possible even before widespread adoption of digital platforms. Bogard (1996, p. 27) explained a “profile as a kind of prior ordering”, essentially a computerised sorting of individuals for the purpose of automated decision-making. According to Bogard (1996) this was the first step to using data for building surveillance. The focus of Bogard (1996) was on the surveillance systems built mainly by both government and private organizations. The review of literature is focused on the profiling of media audiences and the development of advertising technology to profile and target users of digital platforms.

Whitaker (2000, p. 71) wrote about the task of converting data into information and outlined that “information is a product”. The process of gathering data about individuals and creating an information profile against which advertisers can create their target audience has been defined in the academic literature since the early days of the internet. Westin (2000) weigh in on the debate

about the amount of data required to service the requirements of a consumer by providing two sides of the debate. On the one side is the dominant business model: "We must know you to serve you." On the other side is the majority consumer model: "I'll decide what you know about me". Westin frames these two pertinent positions at the early stage of the adoption of the internet and these two positions remain relevant as the debate about consumer data, privacy and advertising surveillance is still raging. Elmer (2003, p. 115) described:

> Surveillance on the Web is a similar means of sorting and categorizing users and is enabled by a similar logic of connectivity, unlike how hyperlinks connect documents. Web surveillance is largely concerned with facilitating reliable and continuous connections between users and (largely) commercial servers.

These were early days of internet adoption and structures enabling web surveillance were taking shape. Scholarship exploring the mechanism and building blocks of this surveillance system was catching up. To be successful in the field of e-commerce and become a serious advertising force, web surveillance needed a system to track an individual across the internet.

Building consumer profiles in the digital world starts with assigning them a unique identifier, through which they can be tracked across platforms. Over the years, the mechanics of tracking a consumer across websites have been studied and documented. Elmer (2003, p. 118) had documented the most prolific identifier that has been around the beginning of the internet as a mass medium: "Cookies essentially provide servers (and their owners) a means of identifying repeat visitors to their websites". After its introduction in December 1994, the simple cookie went on a journey and became a means of giant strides in advertising technology. According to Elmer (2003), these few lines of codes called cookies became the identifier that tracked individuals across the internet. The cookie was transformed into an advertising force by online advertising giant Doubleclick (now Google Marketing Platform), who partnered with multiple websites to drop its cookie on the machine of visitors. In doing so, it created the first detailed profile of an individual without explicitly asking for submission of detailed activity logs by the individuals (Elmer 2003). Wall (2006) defined the cookie and how it is used for surveillance on behalf of individual websites or for an advertising technology company like Doubleclick:

> The primary function of a cookie is to simulate a continuous connection to a website and therefore assist the user to navigate the site with as little obstruction as possible. However, in carrying out its intended purpose it also performs a clandestine and automatic function of transmitting personal information to the web server once a specific web page has been selected. It thereby enables the owners of the website to

> surveil the users in order to obtain important data about them, for example, about the frequency of use of a particular website and what is being accessed on it.

In addition to Elmer, Wall (2006) also discussed the origin of cookies as a simple code tasked to provide personalisation to repeat visitors of a website and how it was converted into a surveillance powerhouse. Individuals are converted into affinities and intents that can be used for advertising products and services. The intent of digital platforms is not only to collect "big data" but also to deploy analytics to convert it into meaningful information. Over the years other identifiers emerged and were used in combination to track individuals across time and digital platforms. McStay (2017) described the use of a double click cookie, Google AdID and Apple's IDFA as the identifiers used to get "intimate" with the consumers and create a profile of users without using personal identifiable information (PII), which is regulated to various degrees across countries. The combination of tracking technology and analytical abilities developed over time combined to form the backbone of ad tech.

This usage of big data for the purpose of surveillance had been described by McStay (2009, p. 101) as "dataveillance" and elaborated as the "use of personalized information for segmentation, clustering and predictive purposes". Advertising-funded media platforms provide free software-based services to users, who in turn consent to share their usage. Van Dijck (2013, p. 31) provides a detailed definition of data in the context of usage of digital platforms:

> Data can be any type of information in a form suitable for use with a computer, for example, text, image, sound, and numbers. Data may also be personal profile facts, such as name, gender, date of birth, and zip code, commonly provided by users following a site's protocol for registering and entering. Metadata contain structured information to describe, explain, and locate information resources or otherwise make it easier to retrieve, use, or manage them.

Another important aspect of data-based surveillance is what Mayer-Schönberger and Cukier (2013, p. 78) classified as "datafication" and described it "to datafy a phenomenon is to put it in a quantified format so it can be tabulated and analysed". McStay (2018, p. 97) wrote about the process of converting people into "data doubles" and mapping this virtual identity to "interests, outlook, beliefs, emotions, age, life-stage, income level, relationship status". Similarly, Szulc (2018, p. 3) described "datafication"; the practice of creating profiles on social networking services (SNS) as "a double imperative to gather as much data as possible and to pinpoint the data to a particular unit". All these definitions illustrate the capabilities of digital platforms to capture the actions and physical movements of a consumer and create profiles that can be used for delivering targeted advertising.

Dataveillance requires behavioural data as exhibited above, but other qualifying data like metadata is crucial for audience profiling as well as targeting. Lyon (2014, p. 3) defined metadata as “data about data” such as the IP address, the identity of the contact, the location of calls or messages, and the duration of the contact. As the definition suggests, metadata is used to add depth to the data collected about individuals and some of the metadata could also be used to identify the individual or device from which it is collected. Mayernik and Acker (2018) described the growth of metadata and the opaque nature of its usage: “As metadata from our digital communication grow in volume and our use of digital tools proliferates, they will be used and aggregated in ways both known and unknown, intended and opportunistic”. They are describing the necessity of metadata for a number of industries to function but view the linking of existing metadata in the offline world with the digital data trails of individuals as something that requires continuous investigation and oversight.

This combination of behavioural data and surrounding metadata is what Glass and Callahan (2014, p. 43) called the “data muscle” of Google and Amazon where more use of their platform and the resultant data makes them stronger. They have used various examples of companies like Dell and Sears to showcase the use of customer data to build strong businesses even when there was no internet or digital data trails. Glass and Callahan (2014, pp. 135–136) talked about how “Netflix flexes its data muscle” by building connections between different pieces of content watched by the same users. The platform then cross promotes its shows to audiences watching the same artist or genre. Digital platforms use historical data to make predictions about the future viewing affinities and are presented to the users as recommendations. Mayer-Schönberger and Cukier (2013) emphasise the need for profitability as the impetus for big technology to dominate “big data”. Flyverbom (2019) wrote about the digital traces that the users leave behind: “Internet companies have access to mind-blowing amounts of data showing everything we do, care about and search for in digital spaces”. Feature engineering is used on “big data” to identify cohorts or “lookalike audiences” desired by advertisers. This unique ability of digital platforms to collect and manipulate large datasets is of critical importance as multitasking on digital devices has created attention fragmentation (see section 1.4.1).

Brasel and Gips (2011) undertook primary research to understand the multitasking between computer and television and found that most of the time was spent on a computer screen with longer gazes as compared to television. The study did not include mobile devices so its implication for today’s mobile-first digital environment is limited but it did show an early trend towards attention fragmenting and multitasking in front of the television set. The attention on television advertising is also undergoing a massive shift and According to Zackon et al (2017), a solo viewer watching television pays attention only to 58% of the ads and attention drops to 25% when there

is a second person and a second screen present. According to research by Zackon et al, a second screen and other activities compete for viewer attention at least 40% of the time. This level of distraction and attention fragmentation makes it important for digital platforms to create accurate profiles of their users and match it to the requirements of advertisers in order to deliver the most relevant message. McStay (2018, pp. 97–98) defined data in the advertising context as "meaningful information used to affect behaviour and experience". The importance of repeat usage and engagement is what digital platforms aim for and understanding the impact of data practices on consumer behaviour.

The fragmentation in traditional media is mirrored in the digital realm as well and a common identifier is essential for creating a complete picture of digital activities at an individual level. Attention fragmentation has ensured that data driven audience profiles become paramount for the survival and success of any digital platform. Skeggs and Yuill (2016 a, p. 391) explored the relation between the strength of respondents' Facebook network and the accuracy with which Facebook algorithms were able to deliver relevant advertising in the users' news feed and found that respondents with "dense and influential" networks received highly customised advertising as compared to respondents with loose, small and less influential networks. Skeggs and Yuill (2016 a, p. 391) attributed this variance to the ability of EdgeRank algorithms to match individual users to a "predictive persona" that is desirable to the advertisers.

The financial success of a digital platform is dependent on retaining and engaging an advertiser-preferred cohort of users and creating accurate affinity groups. Shöenberger (2017, pp. 300–301) discussed the benefits of digital advertisers crunching huge amounts of user data to create relevant advertising but raised questions about the issue of consumer knowledge and control over their usage data. Today almost all of the advertising on digital platforms is delivered via real-time advertising auctions, described by McStay (2017, p. 146) as taking place in the time a web page loads on the browser:

> The process is that:
>
> – a person uses an online device;
>
> – the device requests an ad from an ad server;
>
> – the request is redirected to an ad exchange;
>
> – the opportunity to present an ad is bid for in real-time;
>
> – the winner serves the ad;

– the person sees the ad.

For a real-time auction to work, the profiles have to be created before the fact and have to be refreshed every day to provide value to the advertisers. Tsesis (2018, p. 1596) described the generation of customer profiles based on data collected, joined and stored indefinitely by digital platforms:

> Algorithmically obtained information is disseminated based on models created via artificial intelligence outcomes. Tidbits of information are then tagged to create computer profiles based on algorithmic models. Social media companies keep track of people through internet service provider addresses, data acquired through sales, public records, listservs, sharing websites, and so forth.

Tsesis (2018) highlighted the extent to which the online and offline behavioural and transactional data are merged to create advanced advertising profiles that fuel the real-time auctions described by McStay (2017).

This section developed a perspective on the importance of big data to create accurate audience profiles. Scholarship exploring the importance of digital traces for AFDP and processes used by ad tech to track individuals across time and platforms is reviewed. This section established that audiences are trading their attention and opting in to be observed via their usage data for access to subsidized or “free” use of services delivered through digital platforms. This theoretical construct of exchange does not include informed consent and the voice of the consumer is missing from these discussions. There is a conceptualization of heuristics deployed by everyday users of complex information systems to understand its functioning and is explored in the next section.

2.4.5 Folk Theories

This section explores scholarship about folk theories developed by everyday users to understand complex technologies and to develop a simplified understanding of the world around them.

Toff and Nielsen (2018, p. 637) defined folk theory with regard to usage of digital media as, “the culturally available symbolic resources that people use to make sense of their own media and information practices”. Their focus was on the particular field of media and information as people are intrigued by the processes that create both broadcast as well as digital media platforms. Increased time poverty is one of the reasons people take short cuts in developing a theory to understand a particular issue.

Andrejevic (2013, p. 114) described the complexity that seemingly infinite amount of information on the internet creates for users and calls it a glut of information. Everyday users find it difficult to make sense of such a complex information environment. The invisibility of ad tech and interconnectedness of various elements as described by Skeggs and Yuill (2016b, p. 1367)):"the advertising Facebook targeted, via stories on a user's newsfeed, was drawing not just on data collected on a participant's Facebook page but from all other internet sites they use with a Facebook connection." Skeggs and Yuill were able to decipher this level of complexity by employing a complicated research design using Facebook Application Program Interface (API). Unaided, users are not capable of undertaking such measures to understand customised advertising. The mitigating strategy of the users is pretty similar where they tend to find a knowledge shortcut.

Toff and Nielsen (2018) researched infrequent users of traditional news and identified three folk theories that the respondents relied on to navigate their "information environment". Most of the subjects were content with the news that was being shared on their social feed ("news finds me") or they felt that all the information was easily accessible via a Google search ("the information is out there") and a small minority who found the volume of information too cumbersome to make sense of ("I don't know what to believe"). Toff and Nielsen (2018) chose their respondents from lower socio-economic backgrounds to target news avoiders and brought out the ever-important role of social feeds as the first contact with the information environment. The respondents were satisfied with indirect engagement with "professionally produced journalism" and "distributed discovery" of information useful to them. The folk theory of "news finds me" is an example, wherein news consumers have disengaged from traditional news outlets to optimize their time spent on searching relevant news but are instead satisfied with indirect engagement with the news via their social feeds. They believed that if the news is of importance, they will discover it via their social media feeds.

An early indication of developments in the area where people get the information by asking for it comes from Daniel and Palmer (2007) who reported the future vision of Eric Schmidt, CEO of Google as:

> We are very early in the total information we have within Google. The algorithms will get better and we will get better at personalisation. The goal is to enable Google users to be able to ask the question such as 'What shall I do tomorrow?' and 'What job shall I take?'

In 2007, the growth of the predictive power of algorithms heralded the beginning of a race to get the most accurate information about audiences and the context surrounding their digital

behaviour. Here, Schmidt is predicting the direction where Google search was headed. He predicted that in the next five years, Google's algorithms would be able to give answers to questions that are not even asked. In this future, an algorithm with enough information about an individual is able to create contextual information about the most important answers without actively being asked. This seems to be the foundation on which the folk theory of "news will find me" stands.

This process of the algorithm getting better with incremental data about individual users is described by Mayer-Schönberger and Cukier (2013, p. 201) as "inductive probability" that is crucial for big data analytics, wherein the accuracy of prediction improves with new information being added to the analysis. This is what Eric Schmidt was referring to in 2007, when he predicted the improving quality of algorithmic output with more data being gathered about Google users. Bucher (2017, p.13) used the term "the algorithmic imaginary" to describe how users "imagine, perceive and experience algorithms and what these imaginations make possible". The algorithmic imaginary is people trying to create meaning and modifying their actions to get the best out of their engagement with digital platforms like Facebook. Users of digital platforms may not understand the mechanics of how their feeds are curated but they do understand that their actions have impact and devise folk theories like "news finds me" and act accordingly.

DeVito et al (2018) explored how users form their folk theories about how "algorithmic social media feeds" are populated and found that 89% were a result of participants becoming aware of "endogenous information" on Facebook platforms. Endogenous information is something that participants pick up about how their feeds and the feeds of people around them were getting populated and tried to build folk theories around them. Participants in their study combined learning from observing their Facebook feeds and complemented this information from "exogenous sources" to tailor their posting behaviour. The dominant source of influence was also dependent on the skill level of the participants and participants with lower internet usage skills were dependent on eternal intervention to make sense of Facebook's algorithms.

As these folk theories were malleable, they changed with users getting new information that could be related to frequent changes made to the algorithm with the aim of improving engagement with the digital platform. Low web skills give rise to users being open to external influences to make sense of the invisible layers of code that run beneath the user interface. Folk theories are deployed by both corporations and individuals. Rip (2006) wrote about folk theories in information technologies and identified a folk theory named "wow – yuck pattern" or a "hype disappointment cycle" when new technologies are introduced to the general public. According to Rip (2006, p. 353) this folk theory was institutionalized by Gartner Group and was an attempt to

pass off a folk theory as empirical evidence. According to Rip (2006) this was, at the maximum, an informed guess about how new technologies could fare in public opinion.

Other researchers have deployed folk theories to areas like journalism. Nielsen (2016) tried to understand the popular beliefs about local newspapers and found that these beliefs were largely related to how the views of the respondents aligned with the views held by the newspaper in question. Nielsen (2016) brought in the predisposition of the consumers into the mix of other factors outlined above. Palmer (2019; p 335) further extended the use of folk theories to understand how ordinary people who had been interviewed by journalists felt about journalism and discovered that people thought of journalists as "self-serving" individuals. Eslami et al. (2016) documented the process of development of folk theories about the curation of social media feeds of respondents in a lab experiment and found that multiple respondents arrived at similar folk theories. Ytre-Arne and Moe (2020; p 7-8) studied folk theories of algorithms and identified five folk theories: "(1) Algorithms are confining, (2) Algorithms are practical, (3) Algorithms are reductive, (4) Algorithms are intangible, and (5) Algorithms are exploitative". Subsequent to the analysis of these five folk theories, Ytre-Arne and Moe found that the majority of their respondents were "resigned to the existence of algorithms, but irritated by their workings". Use of folk theory helped Ytre-Arne and Moe to understand the sense of resignation respondents felt with regards to the algorithms and an understanding that the continued usage of digital platforms does not mean a complete endorsement of how algorithms were mediating the interaction.

2.5 Research Gap

As seen in Chapter 1, India is one of the largest markets of mobile internet users; however, the majority of digital platforms and advertising networks popular in India are owned and operated by US-based corporations. Most of the media agencies in India are part of global holding companies so the media planning and buying practices are global in nature. As evidenced throughout the review, there is scant literature about India with regards to smartphone usage and specifically research focussed on user interactions with advertising technology.

At an overall level this thesis reviewed scholarship in the field of media and cultural studies. The literature review was focused on three themes: consumer data experiences, data privacy, and surveillance. These themes had a common thread of issues around informed consent by users of digital platforms. Scholarship has evolved over the last two decades about datafication of individuals and the use of algorithmic analysis for dataveillance. The literature describes the use of relational computerised databases for sorting and profiling people based on their transactional data. The advent of the internet and the rapid adoption of mobile computing gave an impetus to

ad tech to create predictive capabilities and profiles of computers to be traded with advertisers. There is scant literature available about the experiences of Indian digital media audiences and advertising technology. As Indian internet users adopt services offered by digital platforms, they are engaging with the same technology and product features as experienced by the rest of the world.

This chapter reviewed scholarship under the themes of data, privacy and surveillance and identified a research gap. The next chapter describes the methods used in this research. Methods chapter will describe ethical issues related to this research, rationale for research design, sampling and recruitment process, analysis of data and assignments of participants.

Chapter 3 Methods

3.1 Introduction

This chapter will describe research design, ethics, data generation, data collection and analysis. Research questions (refer 1.6.1) drive the methods that were used in this research. Research gap pointed to the importance of voice of the everyday user of digital platforms and their data experiences. Methods were chosen to ensure that the research makes useful contribution to address this research gap. Andrejevic et al (2015, p. 385)) described the data mining systems build by big technology giants such as Google and Facebook as " complex, opaque, ineluctable and unassailable" and describe the possibility of a highly predictive system while postulating the superiority of cultural studies to understand the impact of these systems on everyday humans. This study is designed to understand everyday data experiences of young Indians and marketing practices that convert digital data trails into audience for personalised advertising.

3.2 Research Design

Multiple scholars (Humphreys 2011, Kennedy 2016, Skeggs and Yuill 2016a, DeVito et al (2018), Nguyen and Solomon 2018, Kennedy 2018, Toff and Nielsen 2018, Lupton 2021) have used qualitative interviews to explore data experiences of ordinary users. Their research included data experiences of users across various countries when engaging with internet technology and digital platforms. Rabiee (2004) supported focus group interviews as a method to generate a wide range of themes from participants with a knowledge of the subject. In this research an inductive approach was taken to explore the reasons (why) for Indian youth to engaged with established and upcoming digital platforms and the process of (how) determining the safety of the data and information that was generated by their use of these platforms. As the subject of inquiry for this research was data experiences of internet users and with the huge userbase of mobile internet users in India (see 1.2), recruitment was not expected to be a challenge. As described by Edmonds and Kennedy (2012, p. 112) while using qualitative methods:

> Behaviour is generally studied as it occurs naturally with no manipulation of control. The overarching aim of the qualitative method is to understand or interpret phenomena within the context of the meaning that people express without attempting to infer causation or generalize (i.e., external validity) the results to other individuals or populations.

This research studied participants in their local environment and the participants were not asked to travel to be a part of this research. The interactions were kept short with focus group discussions capped at 60 minutes and in-depth interviews capped at 30 minutes. This time frame helped in keeping the discussion focussed on the key themes and did not place a huge time pressure on the participants.

Work for this research started on 9th August 2019 upon receiving ethics confirmation. Interviews were sought from Industry Experts via LinkedIn and details about the research study were shared with initial contacts for everyday users of mobile internet. At the time of the fieldwork, I was residing in New York and travelled to India for the purpose of conducting this research between 24th August 2019 and 12th September 2019 in the cities of Delhi and Mumbai. As I had travelled to these two cities for the sole purpose of research, there was enough time to reflect on the data that was being generated and this informed the subsequent interviews.

I conducted 1 or 2 interviews per day and majority of the day was dedicated to reviewing the data collected, listening to the recording, updating field notes, reflecting on the data, and monitoring for data saturation and updating field notes. These observations were later coded into observations in NVivo. Reflections after every interview helped me focus on questions that generated discussion and avoid asking complicated questions. If the discussion had moved to customised advertising or audio snooping, questions related to advertising were not asked. This approach helped me in getting into deeper discussions in subsequent interviews. These discussions during focus groups across two cities brought out the folk theories of data gathering that are one of the key findings of this research.

The research was designed to have a second phase of fieldwork, if the first phase did not generate rich data for analysis. In discussions with the supervisors, it was decided that the data collected during this phase of study will be transcribed, coded and analysed to explore the possibility of a second phase of interviews. I was personally involved in recruiting the participants for this research and conducted all the research work including interviewing, transcribing, coding and analysis.

In December 2019 I shifted to Southampton to work on the analysis and initial codes for this phase of interviews were shared with the supervisors during January and March of 2020. In the first round of coding, I had coded 45,000 words, 1945 references and 334 codes. After a couple of reviews, the supervisors and I decided that data collected in the first phase had reached a point of saturation. It is important to note that India was placed under one of the strictest lockdowns in March 2020 due to Covid-19. All regular domestic and international flights were suspended. Further, all international travel was prohibited by the university. This combined with the emerging

Covid-19 situation, it was decided to move ahead with analysis from the data gathered from phase 1 of fieldwork and not attempt a second phase.

Nyumba et al (2018, p. 25) reviewed 170 researches and reported benchmarks for qualitative studies in the field of biodiversity conservation: study size of 52 participants, 10 participants per focus group, 7 focus groups per study and 90 minutes per focus group. Similar, Crouch and McKenzie (2006, p. 493) justify study design using in-depth interviews with small (less than 20) sample size as robust: "The mode of research produces concepts and propositions that have construct validity because they make sense as pivotal points in a matrix where interview yield intersects with pre-existing theoretical knowledge". This research had a total of 53 participants, out of which 7 were Industry Experts and rest of the respondents were everyday users of smartphones. 7 focus group interviews and 12 in-depth interviews were conducted. Generally, the design parameters of this research are in the range observed by Nyumba et al (2018). Rabiee (2004) also suggested that four focus groups are sufficient for a simple research question. As this research had two research questions and a mix of focus group discussions and in-depth interviews, it was targeted to conduct eight focus groups and 10 in-depth interviews in Mumbai and Delhi. As per the recommendation of Nyumba et al (2018), focus group discussions should continue until “theoretical saturation” is achieved.

In this research, this saturation was reached after six focus group discussions and five in-depth interviews. The target was to conduct eight focus group discussions equally divided between male and female groups, 10 in-depth interviews with everyday users and 10 in-depth interviews with industry participants. At the end of the fieldwork, there were a total of seven focus group discussions, five in-depth interviews with everyday users and seven in-depth interviews with industry participants. Depth of data generated, and width of topics discussed were assessed after every 4 to 5 interviews and it was evident that the 60 minute time frame that was decided for focus group interviews was sufficient. My prior experience in conducting qualitative research had been that after 60 to 70 minutes participants in a group of 6 to 8 struggle to continue the discussion and repetition creeps in. This research provided further evidence that for a research instrument that is focussed on 4 key themes (digital advertising and data experiences, business models of popular digital platforms, privacy and self-audit) 60 minutes of discussion is sufficient.

Classification details of participants are shown in Table 1. Figure 1 shows the assignments of industry participants and describes the broad industry they were working in.

Age	18 to 24 years	17
	25 to 34 years	29

City	Delhi	28
	Mumbai	18
Gender	Male	20
	Female	26
Participant Classification	Everyday users	46
	Industry Experts	7

Table 1: Classification of participants

Figure 1: Industry Assignment of Industry Experts

3.3 Ethics

This research was guided by university ethics policies. and all research material, including sampling design, participant information sheet (PIS), informed consent form and questions in the discussion guide were approved by ERGO II (Ethics and Research Governance Online). It was the aim of this research to ensure confidentiality of information provided by the participants. I conducted every aspect of this research including recruiting and interviews. Participants were given an option to opt out of the research at any time. Participant information sheet (PIS) clearly stated that they were voluntary participants and under no obligation to contribute to the

discussions or respond to any question that they were not comfortable answering during the interview. PIS also informed the participants; they could terminate the interview any time. Participants were advised to discuss the topic and 1.) to respect the opinions of fellow participants and not show any verbal or non-verbal disrespect, 2.) that they were free to voice their opinion freely and frankly without the fear of being judged and 3.) that there were no right or wrong answers. In one instance the focus group was cancelled as the participants opted out. I was only able to complete seven out of a targeted eight interviews but respected the participants decision to opt out. All the ethical documents (Ethical Documents) were approved through ERGO II application number 52140.

To maintain confidentiality, I transcribed all the interviews.. Names of respondents were replaced by assigned names as described in PIS. The assigned names were in no particular order to maintain confidentiality of information.

3.4 Sampling and Recruitment

Two set of participants were recruited to answer research question 1 and 2 (RQ1 and RQ2). First set of participants were recruited to answer questions about their data experiences and fell within an age, gender bracket. Second set of participants were recruited based on work experience in the media and communication industry. The first set of participants are referred to as **Everyday Users** and work experience in media and communications industry experience was not a consideration in their recruitment. The second set of participants are referred to as **Industry Participants,** as a minimum of 10 years of work experience in the media and communications industry was the only condition for their recruitment. Industry participants will be referred to as Industry Experts throughout this research.

Everyday Users were recruited through snowballing methods. This method of recruitment is consistent with the ethics approval for this research. The first contact being made with an acquaintance thereafter leading to a referral from their social circle. Final participants were those who qualified as per gender and age groups requirements. There are a lot of sensitivities around caste and class in India, so it was decided to leave them out of recruitment criteria. As the research was not aimed at understanding the implications of caste or class on data experiences of participants, the recruitment was done on the basis of age, gender and residency in Mumbai or Delhi.

All relevant information was provided to all participants through participant information sheets. Participants were given an opportunity to respond within 24 hours and thereafter if agreeable,

convene for the focus groups or interviews. Most of the in-depth interviews were conducted at the homes of the participant or cafés close to their workplace.

The industry participants were recruited from media agencies, media platforms and marketing organizations. This was done to ensure a broader understanding of data collection and activation from the point of view of digital platforms, advertisers, media agencies and data brokers. This is consistent with the validation techniques explained by Golafshani (2003).

All the industry participants I interviewed, have had more than a decade of experience in marketing or media planning media agencies (digital or broadcast media), consumer marketing organizations, and media platforms (broadcast or digital). The key reason for selecting senior media professionals was to understand the kind of shifts that may have taken place in industry practices with regards to marketing. As detailed in the introduction section, accelerated growth in the adoption of digital platforms in India has made academic research, government regulation and consumer data literacy play catchup. It was expected that a senior marketing expert would be able to reflect on the changes in the practices observed in the last ten to twenty years. Messages were sent out to contacts sorted in an ascending order of their last names to minimise any selection bias.

As per the original research design, each gender–age group combination was replicated across both the cities to understand consistency in findings for each demographic. Rabiee (2004) recommends the number of participants in a focus group discussion to be between six and ten and an over recruitment as the drop off rate could be as high as a 25%. In keeping with these recommendations, the target was to recruit six participants for all the focus group discussions. The target was met in six focus groups, while in one focus group, only five participants were able to attend. One focus group had to be cancelled due to unavailability of majority of the participants.

As noted in the introduction, India has one of the youngest populations in the world with 75% of Indians under the age of 35. To reflect the young Indian population, Everyday Users in my research were between the age of 18 and 35. The focus groups were divided into two age brackets of 18 to 24 and 25 to 34, to maintain correspondence in discussions. This was done to ensure that the first age group consisted of respondents who had spent most of their teenage years in a mobile internet world (18 to 24). The second group consisted of participants that had migrated from feature phones with very little or no internet connectivity to smartphones (see section 1.1).

There is a fair degree of gender sensitivity in India. Mixed gender focus groups with participants who are strangers often lead to socially awkward situations in India. I have prior experience of conducting qualitative research in India and was aware of the research practice to conduct focus groups along homogenous gender lines to enable participants to express their views freely and frankly. Thus, male and female focus groups were conducted separately in this research. My gender did not pose any challenge in conducting female only focus groups. As none of the questions were related to sensitive topics, this is an acceptable practice in India. On an average 3,900 words were transcribed and coded for female focus groups as compared to 3,800 words for males focus groups. This illustrates the point that there was no lack of enthusiasm to discuss the issues in female groups.

As this research was about data experiences, it was required that all respondents were active users of smartphones, and there was no restriction placed on the number of hours of usage. This was done to ensure the ease of recruiting participants. In terms of media habits, I interviewed participants who were active weekly users (accessed at least once a week) of Amazon, Facebook, Google Search, Twitter, YouTube and television. To keep the study manageable, top two metros (Mumbai and Delhi) were selected to provide geographical diversity and easy accessibility. Participants were recruited from diverse parts of Mumbai and Delhi to ensure that users from both upper and middle class were represented.

Since all the industry participants that were recruited were my acquaintances, there is a possibility of selection bias. On the other hand, due to my extensive experience, it was possible to easily approach respondents who were at the CXO level with multiple decades of experience in both traditional media marketing and digital marketing. The value of data generated due to the high experience profile of the industry respondents was discussed with the supervisors and it was agreed that talking to direct acquaintances was valuable for this research.

3.5 Data generation

The research was conducted using semi-structured interviews and the same research instrument was used for both the Focus group discussions and in-depth interviews for Everyday Users. A discussion guide for focus group discussions (also used interchangeably with group interviews) and in-depth interviews is provided in Appendix A.

3.5.1 Research Instrument

The research instrument for semi structured interview was a discussion guide with 19 questions. Questions 1 to 3 were ice breakers to get the group to start talking about something that was

easy to answer. The answer to question 10 generally emerged from question 3. There was a priori knowledge that the question 2 about CPG brands brings up issues about trust that could be used to draw the participants into discussions about trust in digital platforms. Questions 4 to 7 were designed to evoke discussions about digital advertising, leading into data and privacy discussions. Question 8 was included to gather data about the knowledge of business model of digital media formats and question 9 is a summary of 4 supplementary questions that were initially included to probe the depth of knowledge about advertising as the mainstay of the most popular digital platforms. As part of reflections between interviews, question 9 was not asked in any of the interviews. Similarly questions 14 was asked in the initial interviews and dropped subsequently. Question 15 and 17 were asked infrequently based on the path the discussions had taken in the interview. During the fieldwork, question 13 and 19 were used together to guide the participants to conduct self-audit of privacy and ad preferences.

The research instrument went through various iterations during the ethics approval process. Initially there were 10 questions to be asked in focus group discussions and depth interviews. These 10 questions were supported by 28 supplementary questions. After discussions with the supervisors during the ethics approval, it was decided to merge everything into 19 main questions that were approved by the ethics committee.

These participants were asked questions related to their everyday usage of the internet with a specific focus on the mobile internet. The discussion then moved towards their attitudes to advertising and privacy. Thereafter, participants explored the data collection by digital platforms and the exchange of such data between the participants and digital platforms. Participants were asked if they felt exploited by companies such as Google, Facebook and Amazon. Participants were also asked to conduct a self-audit of apps on their smartphones, and this was used to validate their claims of active privacy management, to understand their data related experiences. Self-audit of mobile phone apps was used to familiarize the participants with algorithmically generated profiles and availability of privacy controls available in the mobile app. The insights generated via quick audit of their phones were used to compare participants' existing awareness of the data collected by digital platforms, their understanding of digital advertising, their attitudes towards advertising in general and customised advertising in particular.

A different instrument was used for the semi-structured interviews with industry participants. The discussion guide for Industry Experts is in Appendix A and the focus of this instrument was to understand the changing availability of consumer data for media planning and buying in India. These experts were asked to outline the processes and ownership for collecting and deploying digital traces of consumers for the creation of customised advertising. Thereafter, they were

asked about the type of discussions on consumer data privacy and practices that took place within their organisations or with marketing partners. Thereafter, questions regarding the usage data available to them for creating their target audience, consumer purchase journey and the process of converting usage data to advertising signals were asked. They were also asked about the most important data that was available from both broadcast and digital platforms for the purpose of marketing decisions. Their attitudes towards consumer data privacy and its importance in the marketing function and the organization was probed.

The research instrument was administered in English and the discussion happened in a mix of English and Hindi. All interviews were audio recorded and transcribed by me and the recordings were later deleted, and transcripts anonymised. The Hindi to English translation was also done by me. All the transcripts were translated into English to maintain consistency of coding and analysis. I used a technique of regular summation, to be certain that voice of the participants was being captured. These regular summaries were coded as observations in the software. These observations were used for respondent validations and their agreement or disagreement was noted and transcribed.

Participants were able to locate their interactions with ads on Facebook, identify advertisers who have uploaded their data to Facebook and figure out how to turn off personalised advertising. This self-audit of Facebook's advertising preferences gave the participants a sense of empowerment knowing that they could exercise control over advertising experience. Some of the respondents actively managed the visibility of their post, none of them had checked their advertising preferences. The participants were able to locate their advertising preferences from Settings and Privacy menu of the Facebook app. Figure 2: Screenshots of Facebook App illustrates the path taken to arrive at information related to advertisers who use off Facebook activities to target advertising.

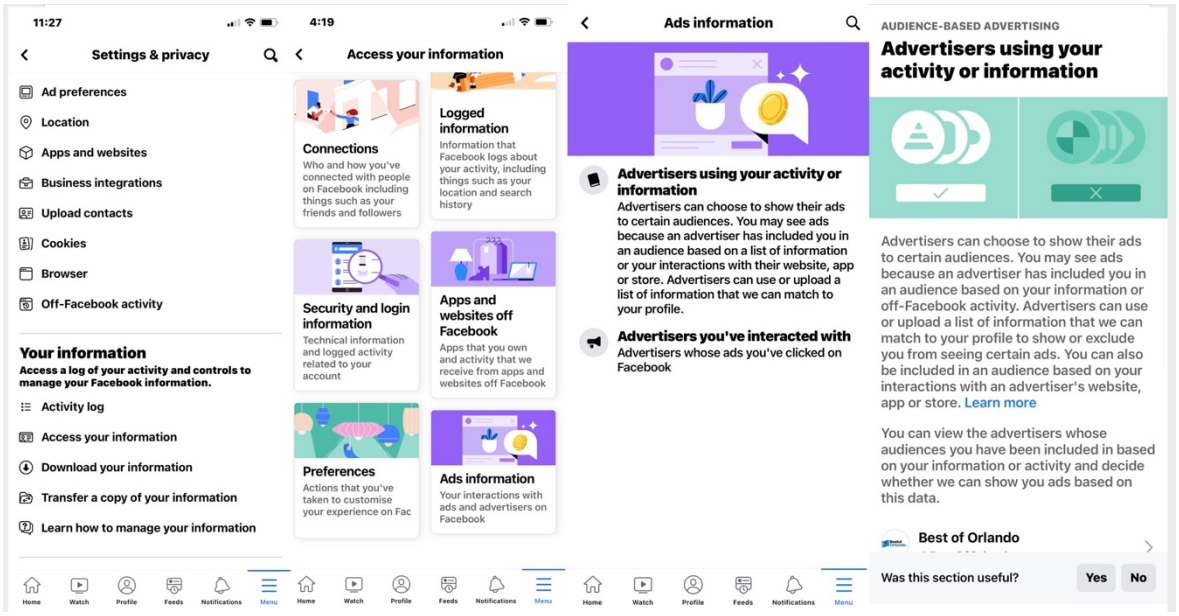

Figure 2: Screenshots of Facebook App

The discussion began with a few sensitizing topics like usage and role of smartphones in their day-to-day life and subsequent topics were taken up from the discussion guide, but the discussion was allowed to flow freely into topics if they naturally occurred in the discussions. Semi-structured interview formats enabled the researcher to let the discussions flow in different directions from a single starting point. The flow of discussions is illustrated in Figure 3, Figure 4 and Figure 5.

This section focussed on the recruitment qualification and methods used to access the participant pool and their contribution to the findings. The next section explores the methods used for analysis of data generated by interviewing the participants.

3.6 Data Analysis

The transcripts were coded using a computer-assisted qualitative data analysis software (CAQDAS) program called NVivo. All the interview transcripts were stored in NVivo and the software was used for coding the transcripts into themes.

Bryman (2016, p. 574) describes Charmaz' approach to grounded theory and coding in three stages: "initial coding", "focused coding" and "theoretical coding" and this was the framework used for coding the transcripts using NVivo. Inductive coding was used to build the final themes from the ground up and no thematic structure was forced on to the initial coding. Initial coding resulted in 65 codes that were combined into 28 themes in the second stage of coding. In the final stage of coding, three key themes emerged that were used for reporting the findings. The discussion of the findings with the literature review was also done using these four themes discovered in the theoretical phase of coding. For the purpose of analysis, the participants were

placed in cohorts identified by the name of the group. As an example, all the participants of focus group 1 were given the prefix of G1 followed by Px. Value of x ranged from 1 to 6 for all groups. This classification helped in the consistent coding of individual participants and in performing individual and group level analysis. Similarly, all in-depth interview participants had a DI prefix followed by gender (MorF), followed by their age. First four participants were from Delhi (refer Appendix C) and the fifth participant of the in-depth interview was from Mumbai. A typical name for a participant was DI1 F23, representing a Female participant from Delhi aged 23. The mapping of age groups for focus group participants is given in Appendix C. Depth interviews with Industry Experts are identifiable by IEx; x has values from 1 to 7.

This naming convention provides the reader with a quick understanding about the person giving the answer. At a single glance it is possible to understand if the respondent was in a focus group or a depth interview, participant's gender, age group and city of residence. This makes it an easier read and every quotation does not have to followed by a description about the person. Industry participants are clearly differentiated by the prefix IE for ease of understanding for the reader. This naming convention removed any mention of caste or religion from the name of the participants and offer an unbiased read on the basis on demographics.

The analysis was grounded in the data and all the themes that emerged after the first round of coding were captured. In the second round of coding, a focused coding approach was used. Themes that emerged from this round of coding guided the researcher to develop an understanding of the key issues being discussed by the respondents. Figure 3 shows the flow of topics during the focus group discussions, through topics that were discussed and how the discussions moved from one topic to another. All the focus groups started with discussions around smartphone usage in most of the focus groups, the participants then moved to discussing advertising. The next set of topics to emerge from the discussions were around the privacy of their data, the privacy red lines that the participants did not want crossed and their awareness about the data that was being collected by digital platforms. As the group discussions progressed, the topics diverged widely and again converged towards customised advertising, surveillance, exchange of data for access, data security and the need for a data audit.

The codebook used for analysis and references coded is attached in Appendix B. A total of 681 references were coded into 65 codes. Some of the key themes that emerged were around data (awareness, control and exchange), attitudes towards advertising, privacy (privacy redlines, important datasets, profiling), Facebook audit, surveillance (safety, dataveillance, trust, folk theories). These were categorised into three main themes of 1) **Data**: This covered all data experiences of participants, 2) **Privacy**: This covered definitions and participant disquiet about the

safety of their data, 3) **Surveillance**: This covered subthemes around advertising surveillance and its impact on participant perceptions of digital advertising.

Self-audit of their advertising preferences was taken up towards the end of focus group discussion and it was undertaken in groups 3 to 7. Participants were not able to access their data or ad preferences on any app other than Facebook. During the self-audit of Facebook ad preferences, participants were able to view the outputs of algorithmic sorting through the interest groups, they were able to audit all the advertisers who shared their data with Facebook for targeting, number of ads targeted at them, number of ads that they interacted with and a host of privacy controls to manage personalised advertising.

The interpretive nature of analysis was able to extract a wide range of data experiences and interactions across participants. As demonstrated in the code book, multiple references were coded for each sub-theme and it helped in validating the research output. Rich data and multiple participants across the two cities (refer Appendix C) provided opportunities to compare key aspects of participant experiences and the researcher was able to reflect upon them from the perspective of wider applicability. One such example was that of folk theories. Folk theories were used to understand participant's process of negotiating with a complex technology environment that had permeated all socio-economic classes in urban India. Folk theory of customised advertising was a theoretical construct that emerged from the analysis, where participants across age groups and cities independently (for the purpose of this research) arrived at the same conclusion - that apps on their phones were listening to them. Some of the other themes that emerged from multiple interviews were: 'fear of missing out' (FOMO) about social interactions and personalised advertising, 'there is no control' (TINC) and 'there is no option' (TINO).

Figure 3: Flow of discussion topics across focus group discussions.

Data from in-depth interviews was analysed simultaneously. Figure 4 exhibits the discussion flow across the five in-depth interviews of Everyday Users and the varying paths the topics took from the same starting point of data awareness. The participants were able to describe their understanding of various topics in depth and as a result, fewer topics were discussed as compared to focus group discussions. However, in-depth interview participants were able to provide a detailed account of their positions with regards to the themes described above.

The next stage of analysis was to identify a representative corpus of extracts from the transcripts that represented the positions of respondents with regards to the top-level themes identified

through theoretical coding. Three themes that emerged from theoretical coding were data, privacy and surveillance. The representative extracts from transcripts are reported in the findings sections and are placed in these four subsections. These findings were analysed in the discussions against contemporary literature to explore similarities and departures to arrive at the theoretical contribution of this research.

Figure 4 Discussion flow of in-depth interviews of Everyday Users

This section described the methods used for this research and the process of data analysis. Qualitative methods were used to collect rich data about participants and their interactions with the complex technological environment of advertising surveillance. The flow of discussions across various interviews was depicted via figures and highlighted the richness of data. Figure 3 and Figure 4 capture different themes across in-depth interviews and focus group discussions.

Figure 6 shows the discussion flow for in-depth interviews with industry experts. The discussion started with changes in user data availability and moved to other themes. Industry expert 1 was able to provide inputs on a wide-ranging topics and data practices for other industries like automobiles. Generally, discussions moved to various sources of data available for media planning

Discussions moved into data ownership and a view of the advertising technology. Industry expert 3 was able to provide a detailed description of data sources and how they were utilised for programmatic advertising. Similarly, Industry expert 5 was able to outline the practices of linking user data across the internet. Industry experts 4 and 6 were able to provide a view of data practices in traditional media organization.

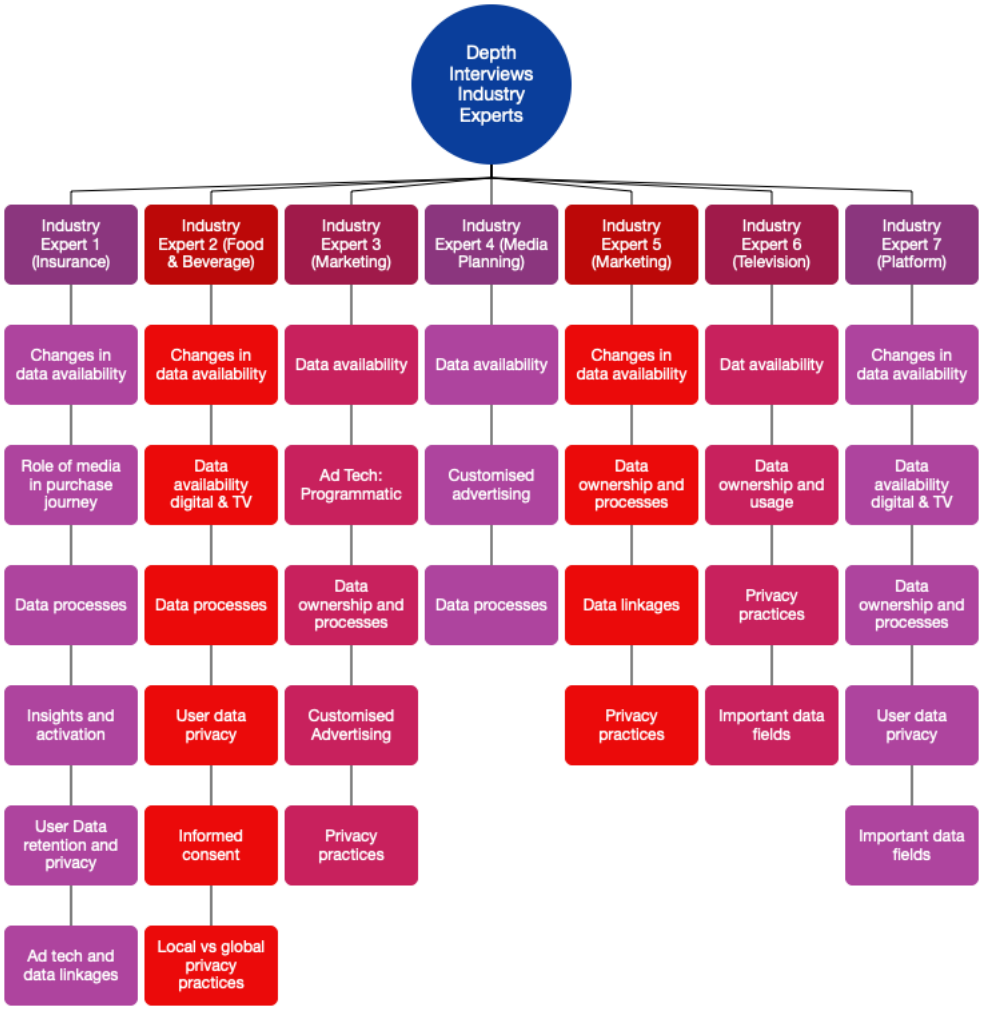

Figure 5: Data flow of in-depth interviews with Industry Experts

Discussions about user privacy were very short except for Industry Experts 1 and 7 who delved on detailed practices inside their organization for managing user data.

3.7 Summary

This chapter highlights the research process starting with the rationale for selection of qualitative interviews as the research method and cities and participants. Two sets of participants were recruited to answer the research questions. First set of participants were recruited to study their data experiences and are referred to as Everyday Users in this research. Second set of participants were recruited to understand data collection and its operationalisation to create customised

audiences for personalised advertising. They are referred to as Industry Experts were recruited based on ten plus years of work experience in the media and communication industry. These two sets are aligned to the two research questions. The approach to data analysis and discussion flow is illustrated through mind maps generated by NVivo.

Self-audit of ad preferences of Facebook is introduced as an innovative method for exposing participants to the quantity and complexity of digital trail data collected by AFDP. My research participants were able to do a self-audit of Facebook app and discovered the output from algorithmic profiling and their assigned interest categories. They were also able to investigate advertisers that target them using their personal information. At the same time participants were able to reflect on this data and possibilities for managing personalisation of digital advertising.

This chapter enlists the approach to developing the research instrument, its impact on data generation and the approach to analysing the data.

Chapter 4 Data

4.1 Introduction

Participants use their smartphones for entertainment, to research information and for online-commerce. These activities leave behind a digital trail data. Mayer and Cukier (2013) and Szulc (2018) call the process of storing and exploiting usage data as datafication. The data profiles thus created are called data doubles by McStay (2018) that expresses the interests, affinities, values, and belief system of the people. These scholars have documented the role of technology in serving the users, creating their data doubles, and building capability to generate information.

First half of this chapter explores the data experiences of Everyday Users of digital platforms and awareness amongst the participants about the type and kind of data that is generated from their online activities. Participants outlined the sense of control over their data and their understanding of data practices of digital platforms. Second half describes the elements that establish trust in a digital platform. These elements will be investigated to understand if there is a gradation in trust amongst different platforms.

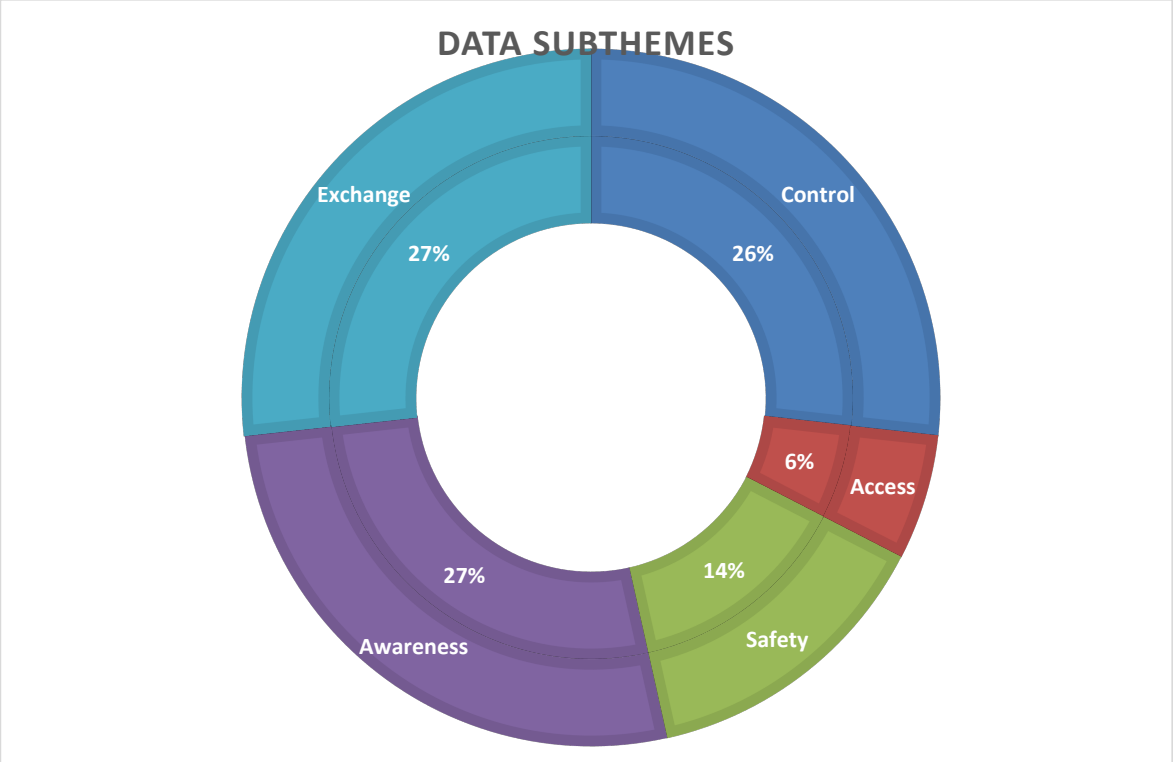

Figure 6: Digital trail data subthemes

Figure 6 shows that participants split their time between three main themes of data awareness, control over their data and the data for access exchange. The percentages reflect the number of

coding references for each sub theme. There was low awareness about the ways and means to access their own data on digital platforms and few concerns were expressed about the safety of their data with digital platforms. The high sense of safety was associated with the level of trust that they had for their favourite digital platforms. There was a trust threshold and once it was crossed participants were willing to share their most personal data with the platform.

4.2 User Data Awareness

The participants of this research can be divided into two groups based on their awareness of the data being captured by digital platforms. The first group of participants held a belief that data entered by them while downloading a mobile app or creating an account is the only information that is available to the platforms. These participants are referred to as Data Aware. A second group that understood the implications of background data collection by the advertising technology. These participants were also aware of algorithms that customised their content and advertising experience. They are referred to as Algorithm Aware. Some of the algorithm aware participants had experience of working in the marketing area of an organization. Others acquired this understanding through their friends who worked in the field of digital communication and through various media reports.

These two classifications will be used to interpret the findings and will also be contrasted with existing literature. DI1 M25 from Delhi details the data that is gathered about him by YouTube and social media platforms:

> YouTube will collect everything that I watch. These days I watch a lot of trailers so my whole feed is filled with trailers. Previously I was preparing for a job, and I was watching puzzles and everything. So, I would get app suggestions like Coursera app, and it will suggest that I should do some courses from there. So, I guess that YouTube collects data from there also, if I am looking at some courses so it will suggest me apps like Coursera, Udemy and Udacity. (DI1 M25 Delhi)

DI1 M25 was able to connect the advertising served to the kind of content that he was consuming on YouTube. The participant provided examples of two distinct types of advertisements that were served: movie trailers and e-learning apps. He believed that the ads were related to the content that was consumed on YouTube. In the past few years, he had gone from preparing for job interviews by watching videos about solving puzzles to watching trailers of upcoming movies after he started working. Bolin and Andersson (2015) bring out the importance of behavioural and transactional data to create a detailed profile of consumers to service advertiser requirement for

insights to target its consumers. DI1 M25 believed that his behavioural data about content consumption was used to serve him personalised ads.

During the in-depth interview, he expressed confidence about his understanding of how his content choice was used to determine his affinities and life stage. He felt that his content-derived life stage affinities thus influenced the type of advertising that he was served. In his student days he was profiled to be suitable as a target consumer for educational and skill enhancement applications, while his changing interests in movie trailers provided signals that he was a target audience for new movie releases. He was also aware of the data those social media sites were collecting and information that was being inferred through that data:

> In social platforms the data that is getting collected is about sentiments, when I like a post, it is garnered from the sentiments that I express about a famous post. Maybe they are doing some text matching, some time back there was something back that they are doing it for politics. In the time of election there are many posts that tilt the sentiments of people. (DI1 M25)

DI1 M25 is algorithm aware and is able to understand his data experiences and can provide complex explanations about content recommendations and personalised advertising about the curation of his social media feed and how his actions on the app were generating data that can be converted into information.

This understanding was shared by other participants who were both data and algorithm aware:

> Somewhere we leave digital footprint, email address, date of birth, all online activity on my phone Insta and Facebook can immediately track it, they can very well track it, have it, store it and share it. You go to Bank Bazaar and put in your phone number and name, and you search for credit card, within half an hour they will give you your complete credit report. I am giving you a general idea, everybody has all the details and you have absolutely no control, everyone is talking about privacy and that there are strict rules, nothing is strict, my activity and my digital media can be tracked by any big company. Whatever privacy are nothing and it has been happening like that only. When I am checking for a shoe on amazon and after half an hour, I can see advertisement of Amazon on Yahoo with the same shoes of Nike and Adidas are visible in the Amazon ad on Yahoo. How is that possible? (G7P5)

G7P5 was able to elaborate on the type and kind of data that is gathered by digital platforms and is also aware of the role of intermediaries in identifying and tracking his behaviour across multiple platforms. This ability to link online behaviour with cross platform identity tracking leading to

personalised advertising is another example of algorithmic awareness. He was also able to chart out the underlying layer of advertising.

Some participants liked personalised recommendations as, platforms could convert the intimate knowledge of their interests into intuitive product features. Innovative and interesting product features were the main drivers for some participants to engage with digital platforms. DI1 M25 described one such feature, where he got a monthly report of the places visited:

> I get this mail from Google. I am not sure but maybe I would have opted in for this mail. There are details about the cities that I have visited, how long I have walked for, how much time I have spent in my vehicle. It also has a funny thing about how much time will it take for the whole revolution of the moon. (DI1 M25)

DI1 M25 was aware of the amount of data that was getting collected and was a voluntary participant in providing the data so that he could get a spatial timeline of his movements throughout the month. Bucher (2017) researched algorithm aware participants to describe data experiences of Facebook users. Even tech savvy participants were jolted by the accuracy of advertising. Unlike Bucher's study, participants in my research were not recruited for this algorithmic awareness, still the expressed a similar level of nuanced understanding.

Similar examples were shared by participants across groups. Some participants had an affirmative approach towards these platforms and their location services were highly valued for being a guiding light in unknown places. G1P6 explained his preference of Google Maps over Siri: "I am a traveller, I go to kind of adverse places in tracking and Google as we were saying, with the help of Google I am able to locate places as well, which Siri is not helping me". Napoli (2014, p. 348 - 349) explains that "algorithms play a central role in producing decision outcomes from these stores of data" and these algorithmic decision making is exposed to the user via product features described above by the participants of my research. If G1P6 prefers one product over another, it reflects his appreciation of decision outcome of Google maps when compared to Siri.

Participants in my research had an affirmative approach towards data linkages between locations generated by GPS in the mobile hardware and its use by digital platforms to provide a service. Humphreys (2011) outlined the frustration that her participants felt when Dodgeball could not pick up their location from the mobile GPS. Similarly, participants in this study had an affirmative approach to product features that were developed using their data generated by their mobile hardware and combined with their platform usage data to provide recommendation.

Prior to this interview, none of the participants had accessed the ad preference section of the Facebook app on their phone. Some of the participants had checked the visibility of their posts

and profile information and a few others had audited which apps were getting data from Facebook:

> Once I had gone to Facebook, on Facebook there is one tab which is called what apps are getting your data, so I saw that the games that we play in our smartphones those also take some data from Facebook. Facebook connected they are. I don't know for some time I might have used some particular game and I have uninstalled those particular games but still those particular games are getting data from Facebook. You have uninstalled from your smartphones, but they are still getting data from Facebook. (DI3 M27)

DI3 M27 who works in Delhi was one of the participants who had gone into his Facebook app to understand which other apps were getting his data and had proactively restricted access to certain games that were still getting his information even after being uninstalled. He was aware of the Cambridge Analytica scandal, desired more control over how and where his data is shared by digital platforms and was data- and algorithm-aware. At the same time, he was not perturbed if he detects that his action on off Facebook sites, trigger customised advertising on Facebook.

> I can understand that Facebook and Instagram and WhatsApp are connected and if something I'm searching on Facebook and those kind of ads are coming on Instagram I can understand. But when I'm searching on make my trip another platform itself and I am getting the ad on Instagram and its noted, so everything is noted. I'm not that worried about it. (DI3M27)

He was not too concerned about the data practices of digital platforms and felt that his data was safe with digital platforms like Facebook and Google. Google Maps was his trusted partner.

> If you see in Google Maps, there is a tab called activity that stores your daily data if your location services are on in your mobile, you can see that in 6-months back on this particular day which location did you travel to they can sell those details as well. the more data you get the ads that you will get can be more focused more customised more targeted towards audience. I feel it's good I feel as compared to random ads Focussed targeted ads are better. (DI3 M27)

He had a favourable view of the data exchange between him and digital platforms. Later in the discussion he expresses his comfort with "focussed ads". His **trust had a gradation**, and he would have more trust for old and established platforms over new platforms. The brand image of the digital platform was also important to establish trust.

A similar trust in Google was expressed by DI4 F27; this trust with older brands came up throughout the interview.

> There is some sort of bias if an app I have been using for a long period of time there is some kind of bias that I have developed, some sort of trust with the app, thinking that nothing is happening till now nothing will happen in the future. But there is a problem with the new app because there has not been any interaction with it. (DI4 F27)

and as Google and Facebook had passed the trust threshold and were brands that she trusted, she felt comfortable with the exchange:

> With Google I have this belief system that they shouldn't be doing anything wrong, they have strong laws in place. So that is why maybe I'm more comfortable using whatever services they are providing, it's brand dependent as well. (DI4 F27)

DI4 F27 used other measures to establish trust in applications by researching reviews posted by other users on Play Store and would not download any app that did not pass this sniff test.

Andrejevic (2007) brought up the use of machines to read and analyse personal information like emails and serving customised advertising based on intelligence gleaned from this exercise. In my research, participants who were aware of this feature were positively inclined to machines reading their information. DI3 M27 enthusiastically described features like a monthly summary of places visited by him and multiple participants described receiving discount offers after searching for certain products. DI1 M25 expressed a similar opinion for the abovementioned feature: “It is very interesting; it is a nice insight about how much I have travelled.” (DI1 M25)

These processes and heuristics provided participants with a sense of trust and once the trust threshold was crossed, participants were very comfortable about sharing their most confidential data with applications. Components of **trust threshold** will be analysed in this chapter.

Providing access to SMS is an example of sensitive data. As most of the applications rely on One Time Password (OTP) for two factor authorisations, they need to have full access to all the SMS to automatically populate it in the required field. G1P4 who works in Delhi feels that it was acceptable to give a mobile app access to his SMS for the convenience of auto population of OTP in the app.

> A good point that I want to raise is nowadays, any application that you install ask for you for three permissions, can I access to your files, access your messages so that I can automatically detect you SMS for OTP and I normally say yes that would even mean that

> whatever SMS I am getting no matter how confidential it is, they have access to it. (G1P4)

Another factor in establishing trust was the fair use of sensitive data like phone numbers, location, email addresses and contact lists. Participants understood that if they wanted home delivery of groceries, they would need to provide the website with their address and phone number to fulfil the request. Fair use of data also pushes an app across the trust threshold.

The requirements to be on popular platforms and to be in the know of the best deals were put forth as the key reasons for continued use of digital platforms without a lot of research by G5P2 who is an early career professional working in Mumbai.

> We say we do not have time but to put it in a very 8 hours in office, three hours we are on Facebook or Instagram in between, because if you open Instagram once you will end up spending 15 minutes on it. So, we cannot say that we do not have time. But if you ask us maybe because we live in the world right now where everything is on Facebook and YouTube and others and that our life cannot move ahead without it. That is why we are there. (G5P2)

G5P2's response sums up this section as she takes responsibility for not spending enough time reading the terms and conditions and provides two key reasons for continually using the digital platforms: 1) **There Is No Option (TINO)** and 2) **Fear of Missing Out (FOMO)**. TINO and FOMO will be explored in the next few sections to establish if the concepts are commonplace across different groups. Pybus et al (2015) discovered that FOMO was prevalent amongst their participants and it motivated the users to continue their patronage of digital platforms.

My research found that participants in India had a good understanding of three of the five domains that were presented by Pangrazio and Selwyn (2019). The degree of comprehension about 1) data identification, 2) data understanding and 3) data reflexivity was high. Participants were aware of the data that they provide when they signed up for any app or free trail on a website. Some participants were able to identify the data that was collected by their devices and apps and how it was used to target them or follow them across the internet. Examples that participants provided were: my activity on a website trigger ads on Facebook or Instagram, location awareness of advertising and contextual awareness of advertising. The understanding of uses of data was high for Everyday Users who were working in the media and communications industry, and some non-industry participants were equally aware.

The examples of deployment of data tactics were rare in this research, though some of the participants used personalised advertising and content recommendations for surveiling their

favourite brands and product categories. In my research participants with data understanding are referred to as data aware and participants with data reflexivity are referred to as algorithm aware.

4.3 User Data Control

This section builds on the participants' awareness of their usage data and the level of control that they have and explores how they exercise control and manage their visibility to advertisers. Throughout the research, some participants believed that they had no control over what data was being collected or how it was shared with advertisers. Some participants felt that they were not able to seek any control over their data from digital platforms and had to consent to all the permissions sought by the digital platforms to access the app. Andrejevic (2014) brought up the issue of unequal balance of power between users and corporations, similarly, Carmi (2020) has outlined issues related to power imbalance between advertiser visibility (resulting control) over users and lack of C2B visibility.

As explained by G5P2 in the previous section, TINO and FOMO are the key drivers behind granting all permissions to the trusted digital apps. G6P4 felt that the quantum of control depends on the organization building the application. "We cannot control it. We can control in the way the company wants. We cannot control our information the way we want to". This sense of **There Is No Control (TINC)** was observed in most of the focus groups. Some of the participants were excited to discover this control after the self-audit of their phone. Other participants were apathetic to the controls as they wanted to receive focussed, customised advertising that was useful information to them.

The issue of lack of control does not seem to impact their usage of the platform like Facebook but the product features that are created by using their data made them uncomfortable:

> I was in Bandra, so it showed six friends near you in Bandra. How? Maybe because I have allowed Facebook to give me information if there are friends nearby. It only found out that I am roaming around in Bandra and there are six people known to you are around in the area. Say Hi! I do not want to say Hi! (G6P4)

G6P5 who is a homemaker living in Mumbai liked using Facebook, but she could not control her usage data to her satisfaction levels, including her location. In the example provided above, she did not want her Facebook friends to know where she was in her city and wanted her location to be private.

The product feature that was trying to put her in touch with other friends in the geo-location was not received favourably by her. The lack of explanation around such product features also added to the sense of resignation. If she is not able to understand how her location information is being collected and is used, she felt resigned to the fact that giving up control is required to access the digital platform. She was very happy with the other functionalities and control over her posts and had a realistic view of Facebook.

Van Dijck and Poell (2013) define the concept of data flow as real time processing of user data and surfacing it as insights about an individual or cohorts. Berry (2011) also describes the movement of data from user to digital platform and vice versa as a continuous flow. The option to broadcast her live location is an excellent example of operationalisation of data streams to provide a product feature to users. In the case of G6P5, this feature was not very useful and highlighted a lack of control over her own data

Some participants understood that the terms and conditions of usage lists out all the permissions that they were providing to digital platforms. There was an understanding that the terms and conditions in End User License Agreement (EULA) were lengthy to read and difficult to understand.

> I was very lazy to read my terms and conditions and hence. I just allowed and I went and used my app now my data is being used and now it is what it is and I cannot do anything about it. (G5P1)

G5P1 discussed time poverty as the reason she did not read the terms and conditions. As a result, she believes her data is shared by digital platforms with third parties. This is the most common theme that emerges from discussions about the understanding of data management by digital platform. Generally, participants who discussed EULA were aware that all the data sharing policies were mentioned in the document, but they did not have the time: **There Is No Time (TINT)** to read them. Participants in Nguyen and Solomon (2018) research in Australia described barriers to understanding data collection practices resulting in a lack of control, similar to the participants of my research.

Another example that emerged during discussion about lack of user control over their data was Truecaller. Truecaller is a mobile app that is used across the globe by mobile phone users to filter calls from people that are not on their contact list. The apps ask for permission to read the contact lists from the phone to provide the service. G5P1 lives in Mumbai and is an earlier career working professional and feels time poor.

> If there is time people like us will search for a new app to replace the app that we already have. Even to read my terms and conditions, I would want them to be simplified. What TrueCaller did was the people who are annoyed with the fact that there are banks and everybody randomly calling them as in spam calls, so they developed an app. (G5P1)

TrueCaller helped people identify callers on a mobile phone that are not in their contact list and filter out unknown and spam callers. Thus, TrueCaller was presented as an example of apps that helped people to protect their privacy and to manage access to their mobile phone.

G5P1 elaborated on how TrueCaller has built a business around providing identity to nameless tele callers and warning people about spam calls. She felt that this was one way people regain some control over their data and can filter out unwanted calls without answering them all the time. She felt that the lengthy terms and conditions were an obfuscation tactic used by digital platforms to get the users to accept them without reading.

Another control mechanism that participants desired was about advertising preferences. Some participants were interested in receiving advertising from favoured advertisers with discount offers that they did not wish to miss. However, there was a sense of helplessness over the lack of control over advertising preferences overall as digital advertising technology was seen as omnipotent:

> There are two things, a) that you actually want to see the ads from these places and that you will miss out on those, b) you don't use the apps and they will use Google to get to us. (G5P1)

G5P1 listed FOMO, TINO and TINC to display her sense of resignation for not being able to engage in meaningful data protection practices. In this case, she felt even if she manages her data permissions on one mobile app, the advertisers will reach her via Google, which she found all-encompassing. This general lack of control was expressed and discussed by some participants across various focus groups and in-depth interviews. Pybus et al (2015) and Pangrazio and Selwyn (2019) also reported FOMO in their research. Pybus et al (2015) also noted that most of their respondents were aware of data gathering but did not want to stop using the social platforms because of FOMO. FOMO and TINO are two of the key themes that emerged during various interviews.

Participants detailed out their lack of control over advertising experiences. G3P5 a homemaker living in Delhi discussed her mechanisms to exercise control over the visibility of their social media posts.

> In Facebook there is an option to public or to friends, so I have selected to friends so it will not be shared publicly. It will be visible to friends only, at times there are some things that we do not want to show at all. (G3P5)

This was the most common form of data management that came up during the research and some of the participants were certain that they can control the visibility of their posts. Some of participants had tried to actively manage other aspects of their data and believed the digital platforms did not offer them this control:

> There is a general feeling that you cannot change your past and it lives with you forever and the same thing is being reflected here that all my past deeds is something that I have no control over, and I can control things going forward. This did change with the "Right to Forget" law that was passed in EU and it also leads to the question of how long does the data that is generated by the action of using the application lives in the database of the digital platform. (DI4 F27)

DI4 F27 was one of the participants who was data-, algorithm- and regulation-aware. In this example mentioned above, DI4 F27 defines **regulation awareness** as the ability to understand the issue of regulation of privacy and recounted examples of countries enforcing it. She presented a well-considered solution to the helplessness expressed by multiple participants about a lack of control over historical data in the digital world. She also raised the issue of the longevity of usage data and a lack of regulatory oversight in India over how long the digital trail data can be stored and used by digital platforms. Low intensity of discussion around regulatory oversight by participants is important to exhibit a general lack of political or public interest in this topic (see 2.3.2).

G1P1 is an early career professional in Delhi was another unconcerned but data and algorithm aware respondent. In the beginning of the focus group, he stated that:

> I am absolutely fine with whatever data they take and believe that whatever I'm doing in my life is absolutely legal, apart from my wife my mother, there's hardly anyone who is bothered about me, so I think if Google cares about what I am doing and where I'm going I'm more than happy with that. (G1P1)

He is comfortable sharing his data and expressed multiple times during the discussion that he is not concerned about the safety of his data with Google. Later during the discussion when asked if the respondents have checked their phones to see the amount of data that was collected by apps on their mobile phone G1P1 stated that he does that regularly.

> What I do personally is I do keep a check on applications that is there on my phone plus the permission that I have given to the applications that are installed. Even though it's trustworthy application I do regularly check the permission that weather I have giving permission to see my contacts I just normally remove or uncheck it. I don't give approvals. (G1P1)

When asked about his previous statement G1P1 was not bothered about Google collecting data as he did not believe that anyone will be interested in his search or viewing history as everything was above board in his life.

He is one of the participants who was conducting a periodic audit of various apps on his phone to check the data sharing permissions. The main reason to do periodic checks was to exercise the knowledge of how permissions could be managed and not worry about digital platforms collecting copious amount of data.

In this section, participants acknowledged that 1) There Is No Control (TINC), 2) There Is No Time (TINT), 3)they felt insignificant as compared to big digital platforms to negotiate for their data rights, 4) managing the visibility of social media posts, watch history and search history were the key control measures that were deployed, and 5) there was no awareness of existing options to manage advertising preferences on digital platforms.

Data exchange was the third most dominant theme discussed in this research and next section explores the understanding of users about the exchange between advertising funded digital platforms, users and advertisers.

4.4 Data Exchange

In the initial stages of interviews, topics like the business model of media platforms and free access to apps were covered. Data exchange was defined as the transaction taking place between participants and the platform for free access to software-based services in exchange for usage data. The question about how participants felt about the data exchange was asked towards the latter half of interview. It usually followed the discussions about data awareness, advertising, data concerns, privacy, and control over their own data. The question around data exchange was asked by explaining the exchange between digital platforms, users and advertisers based on the preceding discussions.

The datafication of individuals had been discussed in literature even before the advent of digital platforms and Gandy (1993) detailed the practice of data gathering, computerised storage and sorting by multiple business to gain a competitive advantage. The main driver for datafied

individuals was the developments in computing speed and accelerated advances in data storage capabilities. Van Dijck (2013) described the data exchange that takes place between consumers and advertisers mediated by digital platforms. Some participants in my research imagined that this exchange was not unique to digital platforms:

> Personally, I can tell you it is not something new. Old example is Television. TRPs are measured by TAM and TAM basically is collecting who is watching what channels at what point in time and basically, we advertise on those channels. If you want a kid watching Cartoon Network. Brands for kids will be advertising on Cartoon Network. It has been happening for ages and we never had a problem with that. (G1P2)

In traditional medium such as television the media audiences are measured through television rating points, and this becomes a currency for trade with advertisers on television channels.

Participants were asked about this exchange of data for access in the middle of the discussion and they were comfortable with exchanging their data for access to digital platforms. Some of the participants did not understand the advertising-based revenue model of the apps and were not even aware that the exchange existed.

Smythe (1981) focussed their attention on audiences and their impact on the commercial success of media products and Meehan (1984) highlighted the importance of third-party audience measurement to produce Television Rating Points (TRPs). TRPs were used as currency to trade media audiences with the advertisers. Caraway (2011) stated that audiences did not participate in the exchange as the exchange happened between advertisers and media platforms based on TRPs. G1P2 described the process of television audience measurement in India and how the TRPs thus generated were used to segment and target the audiences. The discussion in focus group 1 concurred with Meehan and Caraway that the commodity that was traded were TRPs generated by third-party audience measurement companies that were independent of both the media platforms and advertisers.

Some participants were able to describe advertising as the main source of revenue for broadcast and digital media, and there were very few participants who were able to outline the process in such detail as described by G1P2. Einstein (2017) described the traded commodity as the attention of media audiences and described real time bidding (RTB) where there is a real time auction to serve advertising to digital audiences and turned the focus on mediated access to audiences. Industry participants from digital media agencies were able to outline the process of identifying and targeting audiences across the internet, and there was not much discussion about real time bidding.

Fuchs (2012) explored the role of audiences in creating digital media and concluded that the audiences were exploited by digital platforms and were involved in doing unpaid work in both creating and consuming digital media. The research participants when asked about exploitation felt that they were not forced into using a particular platform and made an argument for how the benefits outweighed the risks. None of the participants had negative experiences due to the trade between digital platforms and advertisers, though some of the participants had unfavourable opinions about customised advertising.

The exact question posed was about fairness of exchange and it also triggered a discussion about exploitation. G1P5 an early career professional working in Delhi compared the exchange of data for access to a trade with Satan. Since he relished the fruits of the trade he willingly engaged in this trade. The benefits that were enjoyed were customised recommendations and convenience:

> It’s like selling your soul to the Satan, if I sell my soul to Satan, Satan gives me certain benefits I enjoy. If I have to reach somewhere I use Google Maps. Google gets to know where I go what time I go, what are my consumer habits if I'm searching for the bar if I'm searching for a massage parlour if I'm searching for a channel I'm searching for a movie, anything that I do, so for me Google has made life easy. (G1P5)

G1P5 summed up the affirmative approach that some of the participants had towards sharing their usage data with digital platforms. There was a distinct sense that these platforms have enabled users to become efficient and they thoroughly enjoyed various benefits that accrue because of participating in this exchange.

Whitaker (1999) made an observation about the exchange that consumers make in their day-to-day life while interacting with corporations and government, where data is traded for some benefit or service. This observation implied that increased participation meant increased benefits. My research participants provided two reasons for indulging in this exchange. The first was There Is No Option (TINO), where the services provided by a digital platform was so unique that there was no alternative, so they accepted the terms without reading them. The other topic was There Is No Time (TINT), related to time poverty with regards to investment required in spending time to read all the terms of service. These ad-supported platforms like YouTube, Facebook, Instagram, and Google Search among others. The exchange can be summed up as that an advertising panopticon thrives on participation of audiences and the participants of this research view it as a synergistic relationship rather than an exploitative mechanism.

Search results that incorporated their geospatial data was one of the most sought features, even though it exhibited depth of interconnected data held by the digital platform:

> You can search it in Google nearby XYZ thing. Google knows everything. Let's say CNG pump near me. They know where I am standing, nearby me so what I am trying to tell you is, though our data is not secured, but the services, if we like to take the services, we have to accept that. Irrespective, our data is out there in the market. (G1P5)

Some participants were favourably inclined to sharing their location with Google Search and Google Maps. Affirmative approach towards technology brands guided the overall sentiment of the participants towards Google Maps. This appreciation of technology that solves critical problems for the users extends to relevant advertising.

Google maps also found favourability amongst female respondents living in big cities like Mumbai and Delhi as they felt safe knowing exactly where they were. DI2F23 spoke about the reassuring nature of GPS tracking as:

> I'm not very good with directions so I don't know how will I survive in the city if they were no GPS. I use GPS a lot even when I'm going to a friend's place or meeting or a new restaurant. While giving this survey I feel pretty sad about myself that I am so much dependent on internet. (DI2 F23)

There were instances where the participants felt that the commute estimates of Google Maps were not accurate, but they were appreciative of the service. Some participants in focus group 5 spoke fondly of Google Maps. G5P3 described the experience: "In a place like Bombay even when we are like an hour away, we check on the maps as to how long it will take. That's the hygiene thing that we do every day". Participants had developed their heuristic solution and interpreted the information provided by Google Maps to plan their journeys and account for extra time to get to their offices. They joy of managing the inaccuracies of Google Maps while travelling on Mumbai roads was evident throughout the discussion.

Comfort, convenience, and digital archiving were key reasons for G1P5 to enjoy the exchange, while acknowledging that he may be making a deal with Satan and giving more than he was comfortable doing. In the same focus group, P4 an early career professional gave an example of e-commerce websites to drive home the point about convenience in exchange for sharing data:

> I will take some example let's take Flipkart and Grofers a decade ago or five years ago this was not digitalised whatever I had to buy I would go to a retail store and buy I was not supposed to create an account there and get my mobile number so the convenience I get all my orders are tracked so I can see what I bought three years ago I can also get the invoice if I have to sell it in the platform like OLX again I have to give my phone number and all but these things are helping me. For example, Grofers if I have an

> account If I don't give it I login as a guest nobody would know what I bought, even I would not know. (G1P4)

G1P4 is detailing the efforts required in offline grocery shopping and the amount of time consumed versus the ease of getting groceries delivered at home. G1P4 felt that there was no requirement to provide any data for offline shopping, whereas online shopping saves time and provides convenience. However, online retailers require detailed data about payment, location, and historical preferences to provide the convenience of online shopping.

One of the reasons that users are unable to move away from the most popular digital platforms is the rapid innovation in product features that are deployed regularly. Gurses and van Hoboken (2017) studied the rise of digital platforms to the adoption of an agile philosophy of product management that replaced the old waterfall philosophy where the product launch happens over a long period of time and the end product is tested and released. Adoption of agile methodology of product development enables digital platforms to rapidly launch new features, making it difficult for the everyday user to keep up with the privacy implications. At the same time new features keep the engagement levels high and create an exit barrier for the users.

In the case of online grocery shopping, G1P4 valued comfort and convenience over his data. He felt that the control was in his hands as he could choose to shop as a guest and do away with the requirement of creating a login. Others in the group pointed out the benefits of creating a login, as it would be difficult to manage problems with delivery of the products if they are not registered with the digital platform. This interaction exhibited that it was a shared belief amongst some of the participants that there are clear benefits in sharing data with digital platforms, especially in the case 98of online shopping. In further discussions, it was clarified that the digital platform should only ask for data that is required to deliver the service, for example an address if someone has opted for home delivery.

In another focus group in Mumbai G6P2 a homemaker explained why she wants personalised information about areas of interest in exchange for her expenditure on buying internet data.

> It is a give and take and if we do not give our data, we will not get the knowledge if a new product is coming. If there is a new product coming from J&J or L'Oreal so if we do not see the ad, how will we get to know about it? If we have to watch a movie, we will have to buy a ticket. If we buy a ticket and not watch the movie, how will we know if the movie was good or not? So, in this example the money that we pay for internet data is the ticket we buy for the movie, and the movie is the ads. So, if we have paid for internet data we might as well watch ads to know if they are good or not. (G6P2)

G6P2 wanted to maximise the returns on the money spent on an internet data plan by watching all the ads served on the digital platforms. She wanted to maximize the benefits of internet data by ensuring she did not miss out on information about new products from her favourite brands. There were a few instances where participants had talked about spends on internet data, but it was never described as money spent on watching advertisements and this participant understood that watching ads also involved incurring data charges. For her, it was not only a simple exchange between her usage data for free access to services but also involved the fact that she is spending her time and money on watching the ads.

The participants perceived this transaction to be broader than a simple exchange of data for access. There was an appreciation for the recommendations provided by digital platforms. The participants wanted customised content and information related to their past search history on a digital platform.

> At the same time, it has its pros and cons. So, if at the same time if I'm searching for something, sites already have my data, so I see things similar to the genre I want to watch and I want to read. I think it's convenient I am getting everything in one place and there is a lot of variety from the very same genre. Sometimes I am just scrolling through, and I see those many advertisements. It is annoying. (G2P1)

G2P1 expected the sites which had access to her data to give her relevant recommendations from the genre of content she was interested in and was certain that these sites would be able to deliver such recommendations. This was part of the service expectation that the respondent had of the websites.

On the other hand, advertising while browsing sites was considered annoying and customised advertising was appreciated only when it was related to purposive activities. The annoyance is a case of a very high level of customization where an advertisement is treated as information in a purposive context and annoyance in a casual browsing mode:

> I think it is a fair exchange they need our information to keep a check on our preferences and as per that they show us advertisements and whatever. So that’s this thing so if we need it or not, they have to put it out there. So, I may not need it but the other person may need it, so its personal choice. But I think I am not losing anything, so I think it is fair exchange. (G2P2)

In focus group 2, P1 and P2 were comfortable with the fact that sites have their data and are customising content and advertising based on that. They also accepted that the sites would

deliver customised advertising, but the favourability of advertising depended on the context of the site visited.

This differentiation between purposive surfing, where customisation is acceptable and appreciated versus casual browsing when advertising is viewed as an annoyance was brought up in multiple groups and is a key filter for relevance. **Relevance** was the main attribute that shaped the attitude of participants towards advertising. Some of the participants felt that it is a fair exchange. The exchange as understood by them is parting with data for a desired convenience. This desired convenience is something that was not accessible a few years ago and now it is available at the click of an icon on their mobile phone.

There were instances where the participants lost control over their personal information when they shared their phone numbers to log in to some websites or to explore some lucrative offers. As G1P6 realised, once an advertiser was granted permission to collect Personal Identifiable Information (PII) like a phone number, it was impossible to control the flow of communication from them:

> I searched for shoes, and I added them in the cart and later on they were calling me to buy it. I was getting phone calls. I might have put my phone number to log in so they start calling you to say please buy this and will give you that, we will give you this. So, it was getting very annoying because the number of calls that I was getting was a lot, so it was not nice. It's annoying as If you're in college or in a class they will keep on calling you, even if you cancel, they will call you again. (G1P6)

This issue came up in multiple groups where participants shared their mobile numbers to inquire about offers and were inundated by phone calls. Participants expressed anxiety that once they part with their information, they lose all control over their data and cannot convince the tele marketers to remove them from their database.

In the case of P6, it was the breach of trust by certain websites that were found to be irritating by the participants. A similar sentiment was expressed in other discussions and was detrimental to trust and raised concerns about the safety of data. Trust and data safety go hand in hand as it is not possible for the participants to understand the source of their data leak and it impacts the whole digital ecosystem. This is where the trust threshold becomes central to filter out digital platforms. In this situation older brands from large corporations seem to be preferred by some participants.

Some participants in this study positively described the exchange of data for free access to digital services as a fair exchange and appreciated the use of technology to offer them personalised

content and advertising. There are four key themes that emerge from this section: 1) participants were aware of their data being collected and used for content and advertising customisation, 2) participants appreciated the benefits provided by digital platforms in exchange for their attention and data, 3) participants were concerned when their data was used for relentless follow-up phone calls, and 4) misuse of data by digital platforms raised significant issues of trust and data security.

4.5 Data Safety and Trust

Generally, participants were divided about the safety of their data. At an individual level, participants had a repertoire of frequently used apps that were trusted versus infrequently used apps or new apps that underwent more scrutiny. Participants had a graded response on trust and data safety based on a threshold that was usually based on the popularity and tenure of the digital platform. G1 P3 "My belief is that it is safe otherwise I would not have kept any digital account and I would store it offline. I believe it's safe." G1P3 believed that online data was safe and was not concerned about the safety. To him the mere act of using various digital platforms was a proof that it was safe.

Another reason that came up regularly was a lack of personal experience of malpractice. DI2F23 "I have not had any incident so I cannot pin it down". This is another example where trust was graded and there were some platforms where the participants felt that their data was safe, while other platforms were not trustworthy. Some of the participants of Shklovski et al (2014) expressed a similar sentiment when they were confronted with data collection practices of mobile apps. In my research the biggest influencer of trust was the popularity of a particular platform. Big expressed as many people using it and established platforms described as those having been around for a long time were the most trusted:

> You know we are bothered, but then we see it happening and a large number of population is part of the same concept so we feel that as everyone is doing, so it does not matter. Had it been just to us we might have been bothered much more. Then it’s like as everyone is doing it so it does not matter. (G2P3)

G2P3 a student from Delhi found safety in large numbers with the belief that if no harm is coming to most of the population, it would be safe for her as well. This sense of comfort in large numbers was based on the understanding that established digital brands have a lot to lose if they do not keep their users and their data safe in an event of data breach.

Similarly, G1P4 qualified trusted brands as brands that have been in business for long and this was the simplest definition that was provided by multiple respondents in in-depth interviews and

focus group discussions: “For me the year of establishment really matters if it is a company that is a two-year-old company, or it is a renowned company”. Long established brands that belonged to large corporations were considered safe. Similarly, G1P5 was also confident of the safety of his data with large corporations:

> If there is a top-notch company and the top-notch company is holding my data, they have a lot of security protocol that is followed. People are accountable there. We know we have given this data to this company and the company will also in return know what data secrecy is. (G1P5)

In focus group 7, P3 and P6 who are working professionals living in Mumbai, gave a similar rationale to define big brands and explained their sense of safety while dealing with large corporates. G7P6 equated the size of a brand to safety. “Zomato feels trustworthy, it is a big brand. Any brand that becomes big is trustworthy”. G7P3 made a similar point about big corporations: “I haven't heard that Google has shared their data with anyone, and a person has lost millions of rupees. Because it is also a point that these big brands will keep our data safe”. To be trusted by the participants, mobile applications had to build trust and once they were able to cross a trust threshold, applications could even demand sensitive information regarding financial details and home addresses. It emerged from these discussions that a digital platform owned and promoted by a large corporation would find it easier to cross the barrier of mistrust.

These are five most common qualifications of trust threshold: 1) if the apps were old and used by a large number of people around them, 2) they belonged to brands that they can trust, 3) reviews and ratings on app stores 4) the fairness of data permissions sought by apps and 5) convenience provided by the apps.

Once the trust threshold is crossed some participants used the recommendation algorithms as an inverted surveillance mechanism. There was a sense of entitlement amongst the participants that it is the prerogative of big brands to keep their data safe: “If they are in the market, they will have to provide service and security. A brand is known for security.” (G7P1). Later in the discussion he went on to defend Facebook: “I think when we make Facebook ID, they give all the terms and conditions. Those terms and conditions are there if we look at them in detail.” (G7P1)

G7P1, G7P3, G6P6, G1P5 and G1P4 contributed to an argument that all the information required to understand the rights of users is provided by digital platforms like Facebook. They outlined the need for users to spend time reading and understanding how their data is used to protect their data and privacy. They also felt that the biggest guarantee for the safety of their data is provided by the size of the corporations owning the digital platforms. Participants imagined that it was in

their own interest to safeguard the data, otherwise they stand to lose all their credibility. Participants generally agreed that credibility led to brand trust, to longevity in the market, and to an increasing number of users adopting the platform.

Leading digital platforms such as Google, Facebook and Zomato were referenced throughout the research as being big and around for many years, hence trustworthy. Big was expressed as having millions of customers. There was faith that because they have never been hacked in the past, they will protect consumer data in the future also. Accountability came up as an important factor and it was linked to the responsibility of the brand to protect its customers. The level of discussion suggested that participants had thought about this issue in the past. However, despite of multiple references about Cambridge Analytica some of the participants did not think that it warranted them quitting Facebook. Some participants in focus group 2 did quit Facebook because of data safety concerns, while others were forgiving as the scandal was thought to have happened during US elections and did not present a threat to their data.

boyd and Crawford (2012) described user stories that emerge from digital data trails and my research participants were supportive when these user stories were presented as a product feature. Participants understood that traffic patterns on Google Maps were derived from multiple individuals using the same software and this product feature was viewed positively. The benefit of ease of navigation in a large metro like Mumbai or Delhi was a fair exchange for sharing their location with the application. Similarly, Mayer-Schönberger and Cukier (2013) outlined the usage of digital exhaust by Google to predict future behaviour with participants exhibiting an understanding of some of the most popular features on apps such as Google Search and Google Maps. Participants of Shklovski et al (2014, p. 2351) explained that “the desire to have the application trumped any concerns for data collection”. This behaviour can be expressed as a combination of FOMO and TINO.

Some of these discussions steered towards a comparison between government bodies storing people's data versus private corporations like Google and Facebook storing it. Some participants trusted big, private sector companies to keep their data safe; G1P3 felt that his data was safe with the government as it was managing one of the largest databases like Aadhar. Aadhar is one of the largest biometric databases of citizens in the world, where the government manages the assignment of a universal Aadhar number to all the citizens and it is accepted as a proof of identification for a wide range of government and non-government services. “Government departments like Aadhar. We think it is safe”. G1P3 felt that government departments like Aadhaar have kept the data of Indians safe. In the same focus group, G1P5 was worried about

accountability in government owned or controlled organizations: “In the government sector I feel the accountability is missing”.

G1P2 had similar concerns and further elaborated on the main reason why users trust big technology companies:

> My funda is very simple in this, if my data gets compromised a private company has a lot to lose. The government can get away with it. Today even if a set of people raised that Google or Facebook has compromised my bank account the entire institution of Facebook can be brought down. It will create a domino effect people will start shutting off Facebook immediately. (G1P2)

G1P2 highlighted the strongest pillar of support for trust in big digital platforms; they need to retain the trust of users to maintain their dominance. Technology giants like Google and Facebook have built this trust over the years and participants were advocating these organizations. It was understood that there exists a synergistic relationship between digital platforms and their users. This is the expressed reason for accountability in private organizations as some participants felt that the government cannot be held accountable if they misuse or misplace their data.

Participants of focus group 5 had a detailed discussion where G5P1 felt that Google was the main culprit behind ads following her across the internet. She felt that as Google was linked to all the apps and hence it had the most comprehensive view of her online activities: “I think the root cause of the problem is Google. Because everything that we do is linked with Google”. She had synced her contacts and app logins with one or the other Google product. This convenience prevented her from acting on her convictions that Google was the root cause of all the problems. There was a sense of forced trust and amongst all the digital platforms she felt that her data would be the safest with Google. G5P4 also felt that Google was safer as compared to WhatsApp: “I would rather have my data with Google than WhatsApp.” There was acceptance to the fact that one cannot work and live in today’s society without engaging with digital platforms. Participants like G5P1 and G5P4 felt that they had to choose between a social platform like WhatsApp or Google as their digital vault. This seemed to be a choice that is forced upon these participants and both have chosen a Google product. There were other participants who stored their sensitive information on WhatsApp.

Financial data came across as the most sensitive data type across multiple interviews. Some of the participants expressed concerns about the safety of their financial data if they stored or shared it via WhatsApp; at the same time, they were comfortable storing their identity documents on Gmail. The discussion in focus group 5 was centred around the most sensitive data like bank

account numbers. It was clear to the participants that they cannot keep all the information offline, and they need to carry it with them. G5P3 explained the logic of the choice:

> There is a ranking here but there is a difference. WhatsApp is more of a social thing so you know you are talking to 10 different people and if someone says you know what, if you put your bank details on WhatsApp maybe it goes to their server. Google is more like a personal thing is my thing. The point is like Google is mine. (G5P3)

G5P3 made a distinction between digital trail data and sensitive personal data like financial information. Digital trail data is trusted with some of the largest and most popular platforms. However, when it comes to storing sensitive financial information that requires portability there is a further gradation that occurs. Social media platforms like Facebook and WhatsApp were viewed as public or shared spaces and not considered ideal for storing or sharing sensitive financial information. It was clearly expressed that by their very nature, social platforms are for connection and communication, and one should not put their sensitive personal data there.

Discussions in this section highlighted the key factors that provide confidence to the participants while using digital platforms. These factors were popularity of a platform and the time it has been in existence. Participants felt a sense of safety while sharing their data with digital platforms if they have been around for some time and were big brands used by millions of users. Some participants perceived big private technology companies as having more accountability (see section 4.5, accountability). This perception was based on the feeling that government sector companies could not be trusted as one cannot hold the government accountable.

Some participants trusted government organizations to keep their data safe and talked about Aadhar as an example. There were some participants who had trust issues with the digital platforms. While that did not stop them from using the platforms, they had well-articulated positions about trust and data safety about the internet in general and digital platforms in particular. The key drivers of trust deficit are discussed next.

4.5.1 Trust Deficit

As discussed in 4.1, some of the participants trusted digital platforms; however, there were other participants that distrusted big digital platforms. In focus group number 4 (males aged 25 to 34) there was a long discussion on the trust deficit initiated by G4P4: “If you are going online, you are giving a certificate to all the organization, fine steal my data.” Cambridge Analytica featured prominently as the reason for a very negative image of Facebook and led G4P4 to fear for the safety of their data. Referring to data collection practices of digital platforms G4P4 elaborated

“They are collecting the data and promising that they will not reveal your identity but ultimately what if they reveal your identity and that is the example of the biggest data breach of Facebook.”G4P4 explained his issue of mistrust of digital platform by discussing his fear of a scandal like Cambridge Analytica happening in India and the fact that there were no regulations or safeguards to protect him against either a data breach or the sale of his identity data that is collected by digital platforms. These examples of a data breach and an overall mistrust of digital platforms guided other participants to share their experiences, either gathered first-hand or through news media about PayTM, one of the largest mobiles pay platforms in India:

> There was an instance in which the secretary of PayTM guy, Vijay Shekhar. She had all the information about the users. She was about to give away the information of thousands of people. The information was on a pen drive. (G4P1)

This particular discussion brought out the importance of focus group discussions as, at the start there was a general understanding that data sharing across platforms and websites was beneficial, and it helped them reach better decisions. However, the discussion about Cambridge Analytica brought other issues about data breaches to the fore. The issue with data leaks from Facebook also came up in group 2; G2P3 was sceptical if the data was leaked or was sold by Facebook:

> Convenience is a good thing, but the bad part is when it can be shared with everyone, it is available in the world to everybody. I don’t know the app owners are sharing my data with everybody. Sometime back there was a leak in Facebook where they shared all the private information people had, mobile numbers and everything so there was leak from their end. We do not know if there was a leak or actually, they wanted to sell the information. It's bad obviously. (G2P3)

G4P4 and G2P3 referenced the Cambridge Analytica scandal for their mistrust of Facebook. G2P3 had concerns that were more generic and covered all the apps that she was using on her smartphone. G2P3 highlighted the helplessness that comes with a lack of knowledge about the data practices of digital platforms combined with a sense of resignation that he does not have any redressal available against big digital platforms. In this instance, the big technology companies come across as too big to care. There were concerns that digital platforms such as Facebook are selling their personal phone numbers to make money. These concerns were also expressed by participants who had exhibited awareness that the main revenue source for these platforms is advertising.

As described in section 4.1, most of the participants believed that big companies had an accountability and fiduciary duty towards consumer data. G2P4 explained his reason for continuing to use the digital platforms despite having serious reservations about their ability to protect his data. G2P4 summed up the business model of ad supported internet as:

> If you are going online, you are on a mutual 50:50 agreement that you can do business with my credentials I have no problem and in turn I will be using your services. That is the reality of online world and platforms. (G2P4)

One individual in focus group 2 (P4) had a negative opinion about data exchange and trade, however there was an understanding that this was the only model available, and it boiled down to TINO. Some of the mistrust issues arose from the inability of participants to understand the unseen layer of advertising technology that powers the customization of advertising and its accuracy and timing drove mistrust. Some participants in focus group 1 and focus group 5 worked in media and marketing industry and had a detailed understanding of the working of advertising technology industry. This understanding enabled G5P4 to frame the journey of her digital trail data to information that is used by the realtors to create a customised offer. G5P4 detailed the process:

> I am searching, my search history, my cookies, they are selling something. Some part of what I am doing during the day or where I am going, for example if they know that I visit certain restaurants near Mulund or Bhandup, they will know that I live there. I will automatically get real estate properties, buy Kalpataru in Mulund. How do they know that as Mumbai is a big place? It is because I am searching on Zomato or Google on how to reach there. Which restaurant and supermarket that I visit. So, I am getting geo-targeted ads. (G5P4)

G5P4 described her understanding of how property dealers in Mumbai were able to figure out where she resided, based on the geo-location of restaurants where she ate, and this data combined with her phone number enabled realtors to narrow down properties that would be of interest to her. She worked in the media planning industry and was aware of the world of cookies, geo-location data and surfing history.

Trust also had another dimension, that of usage. It was important for the participants to trust the apps that they used most often. Some participants felt they only needed to manage the security of their data on these apps, as there is nothing to manage on the other apps owing to low or no usage. DI2 F23 felt that she cannot manage the access and privacy for all 100-plus apps on her phone but vocalised that “maybe the apps that I use frequently I should do this because that's

capturing my day-to-day life. Others I might be using once in a year. What kind of data will you generate out of it?"

Steedman et al (2020, p. 825) uncovered "a dialectic and dynamic interplay between trust, scepticism and distrust". Their participants described multiple factors that had an impact on trust in BBC's data practices. Steedman et al also outlines two concerns of their participants that might inhibit them using a personal data store: time and security. The concern about time is similar to TINT that was expressed as one of the reasons the participants of my study do not read the terms and conditions while downloading the apps. The participants in my study had developed detailed processes and had defined trust thresholds that mobile applications had to cross.

One of the solutions prescribed by the participants for overcoming a trust deficit was government regulation. Participants thought of a few government ministries who should be conducting an oversight of data management practices, but there was no certainty if that was happening. Participants across the interviews could not point to a regulator for digital platforms. G1P5 said, "On digital I don't think so. There is no oversight or monitoring". DI3 M27 and G2P3 expressed similar ignorance. DI3 M27 said, "I would say in terms of regulation we're novice, we don't know anything as of now. Both the government and us are novice." G2P3 said, "I think the government should have regulations for the companies to not sell the information." Some participants like G1P5 wanted government regulation of data practices of digital platforms: "It should be government who should oversea the compliance of such things and information". This response from G1P5 was echoed across the two cities but as is evident from the extracts, the responses were short, and this topic was not discussed in great length. This was not surprising as data privacy is not a topic that is discussed in media or a hot button political issue. The fact that India does not have a data privacy law is an example of general apathy towards the topic in the country.

DI2 F23 an early career professional in Delhi was able to present a detailed argument for regulation and why it should be the government's responsibility:

> Accountability bothers me a lot, where do I go? Government has a role to play. Right now, that we're not even flagging this problem but I'm sure that there is a need for data regulation. We are still living in the biblical times where we think that it's harmless. it's not about being harmless or harmful I don't think it's right I think if there is a need for a regulator who sets down cyber laws. we're moving ahead and are already riding the wave and the government is so clueless about it. (DI2 F23)

The question about who should regulate digital platforms did not generate a lot of discussion. Most of the responses were around some form of government regulation with very little detailing about the roles and responsibilities of a possible regulator. There was a sense of resignation that there are no laws to protect either their data or their privacy. DI2 F23 raised the point about users and government not having enough expertise or even knowledge of the challenges presented by the rapid advance of digital platforms and smartphone penetration in the country. She echoed what DI3 M27 referred to, of both the users and government being novices when it came to digital platforms and their use of advertising technology and the scope of regulation that was required or possible.

There was a very high level of awareness about the issues related to Facebook and Cambridge Analytica. Participants theorised that Facebook was selling their data to third parties and generalised it to a folk theory that digital platforms are selling my data to advertisers. The data as described by the participants was their phone numbers. There was a notable level of frustration about aggressive remarketing both online as well as offline through phone calls.

On the topic of a trust deficit, three themes emerged from this section. 1) awareness of the workings of digital advertising (algorithm-aware) but do not trust the advertising technology to be a fiduciary of their data. 2) a belief across interviews that **digital platform are selling their data** for profit. 3) relationship between quantum of usage and trust. Participants could not elaborate on their expectations from a regulator to address the problem of trust deficits reflecting low importance assigned to this topic.

Carmi et al (2020, p. 14) discussed “data thinking” where the respondents were probed about the mechanisms, reasons and modalities of data collection and similar questions were asked of the participants in my research, and those working in the media planning and marketing industries were able to clearly articulate the reasons for data collection. My research discovered that data thinking needs to be divided into two parts where one part would deal with the understanding of participants about data collection practices and another part would explore the folk theories that are developed to understand complex phenomena like advertising technology and its use for generating advertising revenues. Carmi, Yates, Lockley and Pawluczuk (2020, p. 15) outlined “data participation” as a way of engaging proactively with the digital data ecosystem and making efforts to educate peers. Data participation was not prevalent in their respondent base. My research could not find examples where participants were engaging in managing their privacy beyond the visibility of their social media posts.

A participant in focus group 3 narrated how she helped her mother-in-law whose Facebook account got hacked, but this topic did not come up in any other group. There seems to be an

urgent need for the digital platforms to inform their users through in-app promotion about the way advertising preferences can be managed. A quiet release of new privacy features does not help in empowering the users This issue was referenced by participants (DI4 F27, G5P1, DI3 M27) across multiple interviews.

4.6 Conclusion

Historically, traditional media such as broadcasting relied on socio economic classification and age gender demographics to generate audience estimates to trade with advertisers (Bolin and Anderson 2015). Increased adoption of digital platforms offering Software as a Service (SaaS) heralded a shift towards postdemographics audiences (Rogers 2009). Real time data flow (Berry 2011, Van Dijck and Poell 2013, Mayer and Cukier 2014) resulting from exchange of data for access to SaaS between users and digital platforms. The trade of postdemographic audience with advertisers and resultant intimate personalised advertising is the main focus of my research. Pasquale (2015) describes metrification of clickstreams and Sadowski (2020) describes the increasing invisibility of software code or advertising technology that powers AFDP. Real time data flows also provide critical services to the users and participants in my research had a pragmatic approach to the benefits of mobile apps.

My research brings out a graded understanding of data practices by different sets of participants and an affirmative approach towards data sharing in exchange for services. Some of the participants expected this “fair exchange” of data to result in personalised content and advertising that is relevant to their current information and entertainment requirements. This chapter compares the findings of my research with contemporary literature about the role of digital platforms in the life of Everyday Users. This comparison led to the concepts and theoretical contribution of this research to fulfil the research gap identified in the literature review.

Generally, the following reasons were enumerated by the participants for their continued patronage of digital platforms and exchange of information for access to services:

- There is no option (TINO)
- Fear of missing out (FOMO)
- There is no time (TINT)
- There is no control (TINC)
- Comfort in large numbers
- A popular app is used by a cohort of people so should be safe
- Big companies owning popular apps will strongly protect their own reputation

Table 2 Tactics for mitigating data anxieties

S. No	Concept	Context	Significance	Examples
1	There is no option (TINO)	Participants felt that a software service was fulfilling a requirement that no other app could address, hence they would download it without understanding the privacy policies.	Mobile applications with a unique positioning cross the trust threshold by addressing a specific problem of their users.	Uber, Google Maps, Facebook, Instagram, Google search, Pinterest, Naykaa, Amazon, YouTube, Netflix, Hotstar, Sony Liv, PayTM, Flipkart, Wikipedia
2	Fear of missing out (FOMO)	Participants did not want to miss out on social interactions and discount offers so felt demotivated to conduct due diligence about data usage or privacy policies.	It is important for the participants to be able to participate in the social sphere and for the platforms to be at the forefront of enabling these conversations. In the case of e-commerce platforms, it becomes imperative to be able to retarget lapsers through discounts and offers.	Amazon, Facebook, Instagram

3	There is no time (TINT)	Participants felt that digital platforms obfuscate critical information into lengthy terms and conditions and they did not have time to read through them to understand their rights and responsibilities.	Lack of time is a good excuse for managing anxieties about data protection.	N/A
4	There is no control	There is a sense of resignation amongst participants that they have no control over their data trails. This drives apathy towards issues like privacy and advertising surveillance and is also used to mitigate data anxieties.	A sense of lack of control reduces the anxiety of not paying attention to changes to EULAs that are rolled out with updates or enhancement to services.	N/A
5	Comfort in large numbers	Participants expressed hope that platforms used by a major proportion of the	Getting a large userbase becomes critical for digital platforms to remove the barriers of the trust threshold and	Facebook, Instagram, Amazon, Flipkart, Google search

		population would be safe.	enable accelerated adoption.	
6	Corporates will protect their brand equity	Participants had a general understanding that their continued patronage was important for the success of digital platforms and they will ensure protection of their data.	User satisfaction is strongly linked to the security of user data.	Google, Facebook, Amazon, Flipkart

Above mentioned four strategies were aligned with trust threshold as participants only used apps that passed this threshold. As described in the table above, the two key components of trust were the number of people in their cohort using an app and whether it belonged to a big corporation. It was perceived that if a large number of people are using an app, it must be safe. This combined with a sense of security with big corporates, that they will do everything to protect their reputation. Hence, participants were able to justify their usage of popular apps from large corporation to their own selves.

As discussed in (EULA), complex terms and conditions makes it difficult for individuals to understand data practices of apps that they are downloading, and research participants made similar claims about not being able to control their relationship with digital platforms. Every application seeks permission for accessing hardware and software features that creates disquiet for participants. This anxiety feeds into folk theories to understand the complex advertising and targeting systems. These feelings were evident amongst the participants of Shklovski et al (2014) but they too did not report any change in their practices with regards to regularly used apps.

On the other hand, it also requires a mitigation strategy to continue using the services offered by the digital platform. These six tactics (Table 2) were the most dominant that came up spontaneously across multiple interviews. TINO and FOMO were the two most prominent tactics deployed by the participants and were used to mitigate against a sense of resignation whilst offering a rationale to continue using various services as described in Table 2.

Users deployed various coping mechanisms to justify sharing their information with digital platforms:

- Nothing to Hide
- Exchanging data for convenience

Some of these findings are similar to what Pangrazio and Selwyn (2019) and Pybus et al (2015). found with their research, where the respondents were concerned about the data practices of social media platforms but were not willing to quit the platforms due to FOMO. However, this research discovered that users appreciated the digital platforms for the convenience that they provided and found comfort in established brands.

Another marked difference that emerged was a very high level of trust for established big digital brands that was not seen in the study by Pangrazio and Selwyn. In my research trust threshold was a theme that was deployed at different stages of interactions with digital platforms and was referenced throughout the research across gender, age groups and cities. Some participants would reject apps and will not download them from the app store if they do not trust them. The trust threshold for such participants was developed by reading reviews, depended on the permissions sought, the familiarity of the app based on a large userbase, and the immediacy of the requirement forced by a lack of alternatives to the services provided by the app.

Once an app had passed these tests and was installed on the mobile phone, the participants did not bother to check the privacy updates or changes to data sharing practices or explore new controls that were released by major digital platforms over time. The tenure of usage was a big contributor to developing trust over time and older apps tend to be more trusted than new entrants. The trust threshold is also described as a few years of usage without an adverse experience. New applications like Uber that became popular in a relatively short span of time are also trusted as they provide a unique and reliable service to the users. Trust grows when such applications are adopted by a large number of users who are positively predisposed.

India has a poor track record of regulatory oversight for emerging technologies (see 1.3.1) and the absence of a data privacy law in India continues this trend. The topic of regulation did not generate a lot of discussion and debate, however some participants were able to articulate the need for protecting their data in online world (see regulator).

The desire to control how and when their data is used for customised advertising firmly lands in the space of privacy. The next section investigates how users feel about their online privacy and what is the most sensitive data that they want to protect, and their privacy red lines. This section

also explores the understanding of consumer privacy in marketing organizations and the level of discussions that normally happen regarding consumer privacy.

Chapter 5 Privacy

5.1 Introduction

Pasquale's (2015, pp. 32–33) description of advertising technology as "a thousand eyes" was recounted by Industry Experts and was also expressed by participants who felt that the apps that they regularly used had intimate information about their needs and preferences. Industry experts from digital agencies and advertising companies were able to lay out an extensive array of practices that are used to develop postdemographic audiences that Van Dijck (2013) and Bolin and Andersson (2015) described. Activation of postdemographic audiences described above was outlined by Carmi (2020) as the mediated reality presented by software code and algorithms. As described in the previous chapter The participants of my research understood these mechanisms as personalization and some were appreciative of the personalised recommendations. Personalised recommendations like new friends on Facebook or a list of places that a person has visited in a month, were perceived as very beneficial. Meyrowitz (2009) explained how years of television watching has normalised the culture of being on display and participants of my research were comfortable in using the social media platforms, albeit some participants had placed limits to the visibility of their social media posts. Bodle (2016) outlined the power imbalance between internet users and corporations and brings up the issue of inability to the users of digital platforms to access their own information. Carmi (2020) wrote about digital literacy and the power imbalance between consumers and corporates comprised of advertisers and digital platforms.

This chapter explores users' perception of privacy in the digital world, the definition, concerns, management, and their most sensitive information that they wish to protect. The Industry Experts who were interviewed for this research were able to outline the process of using web analytics and advertising cookies to create detailed profiles without seeking user permissions. According to the discussions with Industry Experts, advertisers in India were seeking to adapt to rapidly digitised userbases and aim to construct accurate profiles of their consumers to improve efficiency of their media spends. Privacy thresholds are established by developing an understanding of redlines that participant were not able to cross. These redlines are used to identify sensitivity around certain types of data and personal circumstances that result in privacy compromises. Participants discussed their desire for privacy in the online world and sorted their data into categories to indicate the most sensitive categories. A section is devoted to understanding and enumerating the quality and quantity of discussions that happen inside marketing departments of advertisers and marketers in media agencies in India. The section on

privacy audit describes an innovative methodology to expose users of digital platform to their data exhaust. The audit also showcases the use to algorithmic sorting to create interest and affinity groups based on their digital trail data.

5.2 Privacy Definition

Generally, participants had a grasp on certain aspects of privacy and understood that the key issue that it represented was, control over their personal information. This section explores the definitions put forth by participants and how it impacted their perception of digital platforms. Later, the role of data privacy in marketing decisions will be investigated through interviews with industry experts.

There were multiple definitions of privacy similar to the one expressed by G7P1: “Privacy means that something that only I should be able to access. Others can only see it if I willingly share it with others. Without my permission my personal things are not accessible to others.” G7P1 was talking about individuals, as well as corporate entities who should respect his sense of privacy and wanted to have a say when and how anyone accessed his private information. Personal information and permission were the most important elements of this definition. This is similar to the definition proposed by Westing (2000) and participants across interviews wanted to be in control of their data.

Other more detailed definitions were also given in focus group 7 (P4): “Whatever is my data like my bank details, my Aadhar card, all my digital activities data has to be private. To myself only, if I want to share it with others, only then it should be allowed.” G7P4 felt most protective about assets like his financial assets and his digital usage data The damage to his financial health and his social standing were the most important factors in determining his privacy requirements. Across various discussions, financial information came up as the most important asset that participants wanted to protect.

For others, privacy was more about their personality and the ability to restrict who can intimately know them:

> It’s actually not a monetary value, it is a point of privacy for me. I do not want to be open to everybody. I do not want everybody to know who I am. I am fine with my circle of people to know. I am fine. But any XYZ person to know all of me, without knowing me that's a point of concern for me. (G2P3)

G2P3 expressed her concern about privacy when asked to assign a monetary value to her data. G2P3 defined privacy as access to her true self and that access did not have a monitory value; her

focus was on her data being used to create her emotional persona. She wanted to be understood as a human being and not through information about her that is present in the digital sphere. She wanted complete control over her information.

Some participants provided a lock and key approach to privacy of their data and derived information: “my personal information should be restricted to me. It should not be that everybody has access”. G5P2 wanted to be the only person having they key to her information, and it should be locked to access by anyone else. G5P1 was very conscious about her age and age sensitivity came across in other focus groups as well: “I do not want anyone to know my age”. In focus group 5, G3P5 was sensitive about her age but also wanted control over other aspects of her data:

> We want to hide your own things and do not want to bring them out in public. Like your age, your family members, you have travelled for fun, and you do not want to share this with everyone. My account details should be private. (G3P5)

G3P5’s expectation of privacy covers not only age and relationships but also other information that is captured on social media platforms. This definition covers financial and emotional aspects as well as relationships and brings out the participants’ ability to filter visibility amongst friends and relatives. This control of graded filtering was as important as protecting bank account details. Similar definition was shared by DI4F27 “without my authorisation none of the data should be used or forwarded this is what I understand from privacy”.

Participants defined privacy in terms of control, access, visibility, and selectivity about what information they wanted to share and with whom. The whom could be relatives, friends, or digital platforms. What qualifies as sensitive varied throughout the research, but **financial information came up as the most sensitive data category**. The definitions revealed a tendency towards keeping their world hidden and having control over what they want to share with others. In other words, personal information was under a lock and users desired to hold the key. Participants desired to oversee their privacy decisions and be the decisionmaker about who to share the key with.

The main themes that emerge are: 1) people are aware of the requirement for privacy in the digital world, 2) they have an expectation that their data should not be shared with a third party without their permission, 3) participants were protective about their life stories as well as financial information, 4) privacy had an emotional value as well as a financial aspect to it and 5) there was a general lack of understanding around privacy implications of customised advertising

and other data activations. The next section investigates how participants were managing privacy and if they had the necessary control or understanding to do it to their satisfaction.

5.3 Managing Privacy

In the previous section, the participants expressed an understanding about the concept of privacy. They wanted to keep their sensitive information to themselves and most of them had some red lines that they did not want crossed. In focus group 5, there was a unified response to the question about privacy in the digital world: “there is no privacy”. That response was followed by a sense of resignation related to the topic. Some participants in focus group 5 expressed the reason for their resignation as, “If we do not give permission, we cannot use the app”. Participants understood that if they do not accept all the terms and conditions (defined/sought) by the digital platform, they will not gain access to a particular service or application. There were individual examples of resignation and factors driving it. This sense of resignation is outlined by Phillips (2004, p. 698) as power disparity in negotiations between individuals and corporates.

Some participants from group 5 felt compelled to give consent to TrueCaller for accessing their contact list. “TrueCaller, if you have to use that app, I have to allow it to access my contacts and everything on my phone. I think they ask for it as it becomes their database.” G5P3 was convinced that the benefits offered by TrueCaller outweighed the permissions asked for by the app. In this case too there was a clear understanding that providing their data was a non-negotiable requirement for accessing the app. As G5P3 does not understand how all the data gathered will be used to provide the service, she has an understanding that her contact list will become part of the database of TrueCaller and will be used to identify callers for everyone. A lack of a clear statement on how the user data is collected, stored, and used gives rise to folk theories that will be explored in section 6.4. People do try and put together a vernacular understanding of how their data is carried across different digital platforms and becomes widely available causing privacy concerns.

G3P2 confessed to their role in enabling free circulation of data:

> This flow of data happens because we allow it, they ask us if we can use your media or social media, we do not deny it, but we allow it. Frequently all the people allow it and that is the reason that when you search for one thing it is visible everywhere. (G3P2)

G3P2 described the collective eschewing of responsibility created big data that is interconnected and broadly accessible. G3P2 was one of the few participants who felt that people had some control over their privacy that is usually not exercised.

DI1 M25 expressed similar feelings that control, and moderation were essential and he felt that there are people who do not have control and end up sharing too much information with too many people: “It is a fair exchange, all the information that I am sharing I am pretty OK with that. I keep a control on myself that I have to give so much information only. People don’t have control”. The feeling was that there were enough controls, but people do not make efforts to exercise the control available to them.

Group 2 in Delhi with women aged 18 to 24 had the similar opinion that they were not exploited by digital platforms. Their main source of confidence was popularity of the platforms that they were using. The coping mechanism for these anxieties was that so many people are using digital platforms so there is nothing to worry about. G2P3 called it “mob mentality” and also described her anxieties, “We are bothered, but then we see it happening and a large number of populations is part of the same concept. So, we feel that as everyone is doing so it does not matter”. Participants in focus group 2 spent more time discussing privacy concerns as compared to any other focus groups or in-depth interviews but the time spent on exploitation was not very high in comparison. Most of the discussion on exploitation focussed on explaining why they did not feel exploited.

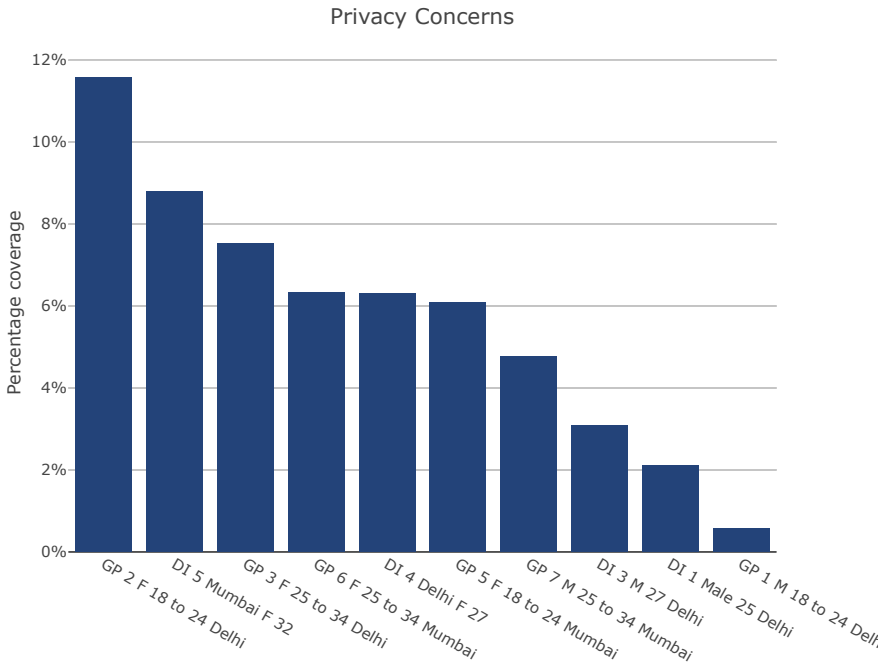

Figure 7: Top groups coded for privacy concerns

Figure 7 shows the coverage for privacy concerns across all the interviews. Group 2 of Female students in Delhi had the maximum coverage. Top 5 interviews by coverage consist of female participants.

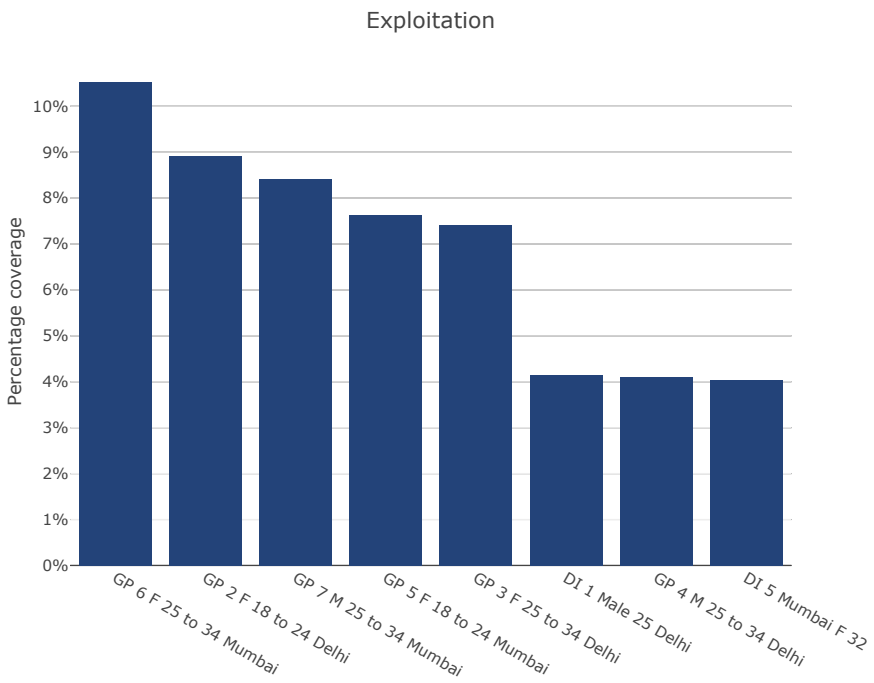

Figure 8: Top groups coded for exploitation

Group 6 spent the maximum amount of time discussing exploitation and some participants were appreciative of platforms with which they engaged on a regular basis:

> What will they do with our data, there is no way we can say that they are exploiting us as we are staying connected with and without Facebook it is impossible to stay connected with so many people. So, it's useful for ourselves. (G6P6)

Being connected was the most affirmative dimension of social media and another participant called it "useful and nourishing":

> At end of the day Facebook is making money because of us I have been using it from 2000. If Facebook exploits us, people will stop using it and millions of earnings that Facebook gets, it will stop. If Facebook has to make money from us it will not exploit us. If that is trustworthy, if there a xyz new platform they might exploit in the beginning, but these platforms are there for 20 years. If they are making money from us and there are so many people on the platform so there would be some kind of security for sure. (G6P4)

G6P4 outlined a synergistic existence with Facebook and both the entities in the relationship lose if the user is exploited or trust is breached. She imagined that user trust will force Facebook to safeguard her data and trust and the same cannot be said about a nascent brand. Her trust threshold was linked to the longevity of the platform and like other participants once an app has

passed the threshold, there was complete trust. Another dimension of affirmation was financial benefits that accrue from engaging with digital platforms, especially for online shopping:

> It is useful for us. Because there is information about discounts that is very useful. There are benefits also for example if you are going to a restaurant for eating out and if you order online from the same restaurant on Zomato, you get a discount. So, there is a benefit. You are saving money. If you are buying mobiles on Flipkart or Amazon so one mobile will be cheaper on Amazon and other will be cheaper on Flipkart. If the brand is trustworthy like Zomato I don't mind sharing my data for the discounts that I get. I will not share my data with a brand I do not trust. (G7P6)

Some participants in focus group 7 believed that digital platforms were not exploiting the users and the exchange of data for targeted advertising had financial incentives. Participants used digital platforms for price comparison and often found financial rewards; they depended on customised advertising campaigns of competing e-commerce organizations to find better prices, deals, and offers for desired brands.

Participants were comfortable sharing data with brands that met their trust threshold. Focus group 3 was amongst the top three groups in discussing exploitation, but some participants had an affirmative approach towards digital platforms. G3P6 listed the reasons for an affirmative approach: “There are plus points, our connectivity has gone up, our awareness has increased. What is happening in the outside world we get to know about it very fast. Else we would not have known about it”. G3P3 had a sense of reciprocity with the digital platforms: “It can be said that they are useful for us and they are also using us”. Other participants in the group (G3P6) emphasised on the act of digital platforms surveilling her environment as useful and did not find them exploitative. On the other hand, G3P3 thought of their relationship as symbiotic and non-exploitative.

Some participants across focus groups 2, 3, 6 and 7 outlined the centrality of digital platforms to their daily routines and espoused an affirmative approach towards them. They described being empowered by their online experiences and noted benefits across the following dimensions: 1) the social connectivity offered by Facebook, 2) financial rewards from price comparison that triggers personalised advertising, 3) surveillance of their environment, 4) a symbiotic relationship built on longevity and trust and 5) safety in large numbers.

Generally, participants felt powerless to negotiate terms with technology platforms. Some participants took proactive action to manage the visibility and reach of their social media feeds:

> The apps give us the liberty to limit our audiences. We can custom them and we can decide who can watch it and who cannot. Sometimes it's up to us what we want to share and with whom we want to share. (G2P1)

G2P1 was aware and felt accountable for the visibility of her social media posts. She also delved into the importance of taking responsibility to make use of the tools found within apps to manage privacy proactively. This is an elaborate understanding of what can be done to safeguard her privacy. In focus group 2, there was another respondent who gave specific examples of how social media apps provided options to manage the visibility of their posts and photographs: "Snapchat, Instagram provide this facility that only your friends can see your photos and everything otherwise it is hidden." (G2P3)

Limiting the visibility of their social media posts was the most common privacy management measure throughout the research:

> I have done it on Facebook I have such kind of privacy on my photos that even my friends cannot comment on my photos. Because I do not want them to comment, and I do not want others to download a particular photo. (G2P5)

Apart from details like phone numbers, age, life stories, relationships, and financial information, photos came up as very **sensitive data** that participants wanted to protect. In terms of managing their privacy on social media, G2P5 has undertaken measures to safeguard her photos and her emotional wellbeing where she had restricted even her friends from commenting on them. Awareness of the ability to manage privacy in the digital world is directly linked to the concerns that participants had about what happens to their data after it has been stored by a digital platform:

> The issue with privacy is that we only give permission for the data to be shared with the platforms, there could be misuse of photos and our information but these days people are open-minded so they allow it. (G3P2)

G3P2 talked about the lack of understanding of the storage and usage of her data after it has been shared with a digital platform. The concern about privacy originated from the lack of awareness on her part and the lack of clear communications on the part of the digital platform.

According to the literature, the key measure of success for digital platforms is their ability to identify, classify and sort their audiences and create audiences for advertisers. The audiences thus generated are enriched by combining data from other sources and successful digital platforms are

able to mitigate the privacy worries of their users while trading them as commodities with advertisers.

In some of the countervailance activities imagined by Raley (2013), to remain anonymous is covered by the range of privacy centric initiatives launched by Apple, starting with the 2017 launch of the Intelligent Tracking Pretension. Since then, Apple has been setting the privacy standards for other mobile operating systems to follow. Safari blocked third party cookies in 2018 and in iOS 15 Apple will block the websites that track IP addresses on Safari by offering Apple Private Relay.

Solove (2004), Shklovski et al (2014) and Shöenberger (2017) agree that there exists a privacy paradox where consumers express concerns about their privacy but do not choose digital service providers on the basis of their privacy policies. According to the discussions in various interviews, the participants recalled a system of evaluating digital platforms and their safety before installing them on their mobile phone through app store reviews, and popularity amongst their cohorts and media reports. However, some of participants had provided their contact details on websites in exchange for free trials or offers. They expressed an opinion that a phone number was not sensitive information.

Algorithmic recommendations exposing personal preferences to family members owing to mobiles being a social device was another consequence of the inability to manage privacy settings. These recommendations could potentially reveal personal surfing behaviours and preferences, to other users of the participant's mobile phone. These preferences could be exposed to other people through customised advertising, content recommendations, customised screens, auto complete, watch history, browsing history and favourites.

There is rich literature available about how digital platforms have developed what Carmi (2020) describes as backend and Turow et al (2015) call data enrichment that forms the backbone of advertising technology. Literature also unveils the point that backend linking of data is where the consumer loses their privacy, but advertising technology is not the prime concern for the consumer. Participants were also willing to put up with telecalling if it was not outright harassment. Participants were concerned about their financial security and were not confident about how to navigate it. They were also concerned about their social safety in terms of control over their personal information related to travel, photographs and their most intimate thoughts coming out in the open to their friends, family, or the broader public

5.4 Privacy Audit

Over the years mobile apps have increasingly provided visibility and privacy control to users. My research tried to determine the ease of locating these resources and the explore the propensity of users to exert control over their advertising experience. To enable the participants to arrive at an informed decision about the ability to manage their data, they were asked to conduct a **self-audit** of their mobile phone and try and locate ad preferences of any app on their phone. Some of the participants were able to locate these preferences on the Facebook app.

After spending a few minutes looking through the available information about themselves, the participants had different reactions to the amount of information that was getting collected. They were able to browse through their advertising profiles, interest cohorts that they were assigned to, the list of advertisers that had uploaded their data to Facebook, controls over customised advertising and a list of advertising exposures in the past. This presented a dilemma for participants about whether they wanted to invest time in managing their data privacy:

> They have just found a loophole of exploiting and making us feel bad about the fact that we have given you permission of doing it because we are lazy as a generation, and we have not read anything and we have just ok'ed it and have just gone ahead with our lives. I still feel it is they the companies or the app builders should control it at least. They are but they need to be very vocal about it. So, for instance the exercise that we did for Facebook, none of us were aware of the fact that we are actually giving them the permission to do that. Now I know that I have a choice and I am willingly not going to do it probably for some of them. I did not know that I had that choice. (G5P1)

G5P1 is reacting to her newfound control over her Facebook feed after auditing her Facebook account and discovering that she has visibility over all the data that Facebook has and the profiles generated after processing the data. She makes a very important point that digital platforms like Facebook are making huge strides in giving users great visibility and control over their privacy; however, these new features that are added to the mobile application are not advertised and there is no in-app navigation assistance available to aid the discovery of such controls. This is in sharp contrast to how new features that enhance engagement are promoted inside these apps.

Similar to DI4 F27, DI3 M27 also highlighted the issue of time poverty:

> There should be a privacy setting across platforms you don't need to go to each and every platform and change your privacy settings they should be uniform platform which will access all the sites all the platforms and give you a uniform access. (DI3 M27)

DI3 M27 wanted a simpler solution of managing privacy across multiple sites and apps on the smartphone. He wanted a simplistic way to manage settings across the entire digital world via a centralised system. This exercise outlined the challenges of scaling the process of **proactive privacy controls** even when easy to use controls are available.

Hundreds of apps on smartphones and a larger number of websites that are visited regularly make it a difficult proposition to actively manage privacy. There was a general sense of resignation amongst the participants that they were in no position to manage their privacy when it involved negotiating with large corporates owning the hugely popular digital platforms. However, the participants were aware of their options to manage the reach and visibility of their social media feeds. Visibility was the most popular privacy management intervention undertaken by the participants in my research. After the participants discovered the availability of options to control their data and advertising preferences, they still felt constrained by their ability to actively manage their privacy due to their extensive digital footprint.

5.4.1 Self Audit as Research Methodology

Participants in this research were able to quickly detect the advertising preferences in their Facebook apps and were surprised at the complete visibility offered by the app. Participants discovered that they could disable ad personalization, browse through a list of advertisers who had uploaded their data to Facebook databases for the purpose of targeting, get a list of advertisers that they have engaged in the recent past and profiles created by Facebook on the basis of their activity on the app.

None of the participants had explored that part of the app but within a matter of minutes they were able to understand this information and could reflect on its implications. “I was not aware that there is so much details that are there with Facebook like ad preferences and there is a possibility of access your information and activities, account ownership I have never read about it”. This is how G3P6 reacted she finished the audit of her Facebook ad preferences. This is how some of the other participants reacted at discovering the visibility and controls that were available to them. The audit also highlighted how people could use these controls to preserve their privacy in India where a mobile phone is not a personal device but can be used by everyone in the household. G4P5 thought of the self audit as a learning exercise to manage their ad preference to disable ad personalization.

> At times there are some content that you do not want others to know about. So if there is some other content, at times my wife or my daughter picks up my phone. I do not

> want them to know or see ads related to that. If I can control it, I would want to control it. (G4P5)

There were some participants who wanted to receive advertising from their favourite brands, and they listed out advertisers that they liked to interact with. G5P1 professed her love for discount offers and decided not to hide personalised ads.

> All these 40 plus advertisers or companies that are advertising it is obviously something that you wanted to see. Air B&B I would go to Air B&B if they are advertising. If I want to eat from Box 8 food, I would obviously want that and Nature's Basket also I want. Consciously I will not say that hide these ads. I am living on offers. (G5P1)

There was a third reaction to the audit of ad preferences in the Facebook app. G5P5 expressed surprise and intrigue at the motivation of Facebook to share this information.

> Why would they give up this kind of information to me? They are holding us responsible; I might not know that this was one of the feature. They are not informing me that this is one way to know where to look. (G5P5)

G5P5 also highlights the lack of in app navigation pointing the users to the ad preferences. G7P6 had a similar observation.

> You can see here all the advertisers that are using your data. We did not know that we had access to his information. Facebook has made it clear that you can see which advertisers are using your data. (G7P6)

These reactions can be summed up as follows: 1) there is extensive information about advertisers on Facebook in the app, 2) there is no in app navigation that guides the user to these controls, 3) participants found these controls empowering and 4) users need proactive communication from digital platforms about privacy controls that are buried deep in the unvisited sections of the app.

Participants were not able to discover this level of detail on YouTube and the control that they could exercise was on their C2C visibility where they could restrict the views of their videos or hide their watch history.

A preference of participants for customised messages in interested product categories was uncovered during the audit of advertising and privacy preferences. One set of participants disabled the functionality of customised advertising and other set decided to enable customised advertising. The second set of participants believed that Facebook had successfully identified their interest categories and the participants were keen to receive personalised advertising as they did

not wish to miss out (FOMO). The second set of participants acted contrary to Bodle (2016) who expected that users would be opposed to customization, as they wanted information and advertising to be customised for categories that they were in interested in.

My research was able to outline a methodological approach of a self-audit of applications to discover privacy control mechanisms that addresses the concerns of Bodle (2016) and Carmi (2020). This relates to the approach of a self-audit as described above in which users use tools already at their disposal to discover the state of privacy control of individual apps. This approach has implications for policy makers to determine the minimum level of controls to be provided through regulatory protections. These tools would be a beneficial addition to the trust threshold that needs to be crossed before the users start engaging with them.

In similar research about data awareness, Pybus et al (2015) had to develop apps to mimic the data gathering practices of digital platforms and develop external tools that could be used by researchers to create an understanding of the type and kind of data generated by internet users and collected by the digital platforms. To explain the magnitude of data collection to their participants, Shklovski et al (2014) required them to download new apps. In my research, the participants without any prior training were able to access their data on Facebook and reflect on its implications.

Reflections included participants wondering why Facebook is openly sharing so much information with its users, participants wanting to check the advertising settings of other family members to see how they are profiled and what are their interests as categorised by Facebook, disabling personalised advertising to feeling satisfied or mocked the selection of categories in which they were placed and looking forward to more personalised advertising.

5.5 Privacy Concerns

The participants were able to express their concerns, about content recommendations and customised advertisements related to their search history on Amazon and Google. Some of the participants had an affirmative approach to advertising that was relevant to their search and interest, but some participants were concerned about the privacy aspect of customization:

> To some extent it is good and to some extent it is bad. But we are using the location services again and again. So, we do not have a problem to this. But from a data privacy point of view, we are secured to some extent and not to some extent. (G4P1)

G4P1 understands the benefit of digital platforms using his location to provide customised advertising or information; however, he does worry about the security of the usage from a data

privacy angle. This debate of convenience versus privacy comes up across many discussions: “There is that option of favourites in the browser, so the browser knows that what we search the most. I think that is also an invasion into our privacy but it's very convenient”. (G2P1)

This is one of the trade-offs that the participants made for enjoying the convenience of autocomplete and autofill. In this example, the convenience of easily finding favourite website on a browser is weighed against the permission to store browsing history. In India, such permissions also have a social aspect attached to them as mobiles can be shared by various family members. Auto fill suggestions, history of web searches, watch history, content recommendation based on what was last watched was considered as breach of their privacy as other people in the household may use their mobile device. G2P4 was not comfortable with product features of digital apps revealing her browsing behaviour:

> Sometimes you look up for a word or something. Sometimes you just wanted the meaning of the word or know about it and then you want to end it but then sometimes it's like Google is repeating it again and again. (G2P4)

G2P4 reiterated that a mobile phone is not a personal device and can be accessed by a lot of people in the family. Some other participants in focus group 2 shared a similar example. The overwhelming sentiment was to find a way to stop very high customization. This sentiment was repeated across different focus groups for both men and women. They do not want others who have access to their phones to know their surfing, viewing, or searching habits. The highly efficient recommendation engines made this a very difficult task.

Mayer (2009) offered a pragmatic approach to the data sharing required for accessing the services offered by digital platforms and wrote about management of privacy. Some of my research participants agreed with Mayer's approach to data sharing with digital platforms and had favourable opinions of platforms that provided them services like Google Search; social networking (Facebook, Instagram, Twitter); location-based services like maps, driving directions, journey time; ride sharing services like Uber and Ola; entertainment like YouTube, Hotstar, news, current affairs, and financial services.

Privacy concerns were most acute around financial data being leaked:

> It is not ill-treating us in any way it's just the phone call that we are getting from the marketers. So, what I am trying to say that there is no harm till the time our bank details are not being disclosed we are fine. (G1P5)

G1P5 was comfortable with tele callers getting his information and would tolerate random phone calls as it did him no harm, but financial data and its safety was of paramount importance:

> Convenience is a good thing, but the bad part is when it can be shared with everyone, it is available in the world to everybody. I don't know the app owners are sharing my data with everybody. Sometime back there was a leak in Facebook where they shared all the private information people had, mobile numbers and everything so there was leak from their end. We do not know if there was a leak or actually, they wanted to sell the information. It's bad obviously. (G2P3)

G2P3 displays a distrust of Facebook and Cambridge Analytica was very fresh in her mind and she harboured a suspicion whether it was a data leak or that Facebook makes money by selling her personal information.

Gandy (1993), Phillips (2004) and Bodle (2016) have described the unequal power distribution between corporates and consumers and some of my participants felt a sense of helplessness when they were not able to control calls from tele callers. Beer (2009) had emphasised on the need to understand the way the data trails created by the users of web 2.0 services were being used to create recommendations and algorithm making decisions in all walks of life. In my research it was observed that on one hand digital platforms are using the participant's data exhaust to gain knowledge about their circumstances; on the other hand, some participants are using the content and advertising recommendations to gain advanced knowledge about product categories of interest.

The data gathering practices of TrueCaller came up in discussions in focus group 4 when the discussion veered towards the tracking capabilities of digital platforms. Participants in the group suggested that:

> It takes all your contacts on its servers and post that it syncs it with other contacts and maps all the phone numbers to individuals and identifies one person to a phone number. So, if a person X is in multiple contact lists with the same number, it is able to identify that this is the name of the person. (G4P4)

The understanding that TrueCaller uses details from the contacts of individual users to build their database came up spontaneously in focus group 4 and focus group 5 where G5P3 made a similar assertion. "As a matter of fact, TrueCaller, if you have to use that app, I have to allow it to access my contacts and everything on my phone. I think they ask for it as it becomes their database". G5P5 was of the opinion that their data is sold to third parties by TrueCaller:

> I am guessing it becomes their database. Sometimes I get a call from a number and I ask them how did they get my number? Some random bank for some debit card or credit card query. I don't know where did you get my number from? (G5P5)

She was pretty certain that TrueCaller pools phone numbers from contact lists of all users to create their superset, that enables the company to identify where a particular phone is originating. She also believed that these phone numbers are also sold to third party businesses. Though some of the participants in other groups took responsibility for sharing their telephone numbers for accessing a service or getting a deal or a quote. Most of the respondents expressed an opinion that TrueCaller or their telecom operator was selling their phone numbers.

> Taking TrueCaller as an example the database that you could get is my first name and my number. Beyond that where do you get my other details. Like my network provider He will have all my details my address and my this and that. So, when I get a call similarly except from a bank, they will have most details. You live in Andheri and your number is this and your last name is this. So, when I put it on TrueCaller, I just had my initial like my first name and my number. How did you get everything? So, When I am giving you some part of my data. (G5P2)

The participant feels that not only is TrueCaller combining her data with other users to get a universal directory of names and phone numbers, but also joining other information to create a more detailed profile.

Another commonplace assertion was about telecom operators selling their data and was directly linked to unsolicited phone calls and text messages received by participants. There was a detailed discussion on the volume of unsolicited calls that users get. One of the most common conclusions was that the telecom operators are selling their mobile phone numbers to various businesses.

> Our network providers are such that they are providing all our information to others. They are just selling data. I am a registered user of Airtel and Airtel is selling all my data to banks and companies like home loans property sellers. (G5P4)

These discussions about data safety led to a discussion about Facebook asking the Indian government to link all Facebook accounts with the national identity number called Aadhar. This came up in focus group 4. G4P1 discussed the request where “Facebook wants the government to dictate that all the Facebook accounts should be linked with the Aadhaar card”.

Westin (2000) and Solove (2004) described privacy as the right of an individual to, control what aspects of their life are exposed to the general public and to selectively disclose personal

information. The privacy definitions provided by participants were in agreement and desired complete control over who can view what information about them. One of the most managed aspects of privacy was visibility of photos on their social media feed. James (2009) studied the American youth and discovered a total disregard for privacy control provided by the social media networks and an apathy towards safeguarding personal information. The participants in this research gave a lot of importance to the visibility of their personal information in the public sphere and more specifically to their social media feeds. Most of the participants had taken measures to protect their information on social media networks. Some early career participants were apathetic towards the dangers of publicly sharing their information, as they felt that they did not have much to lose as they were just beginning their life and did not have a previous adverse experience.

There was general concern about the willingness and ability of various digital platforms to safeguard their data. The opinion was in response to other participants weighing the convenience of using digital platforms with the privacy of their data. There were others in focus group 2 who had removed most of their information from Facebook. In focus group 2, Instagram and Snapchat were considered more favourable as compared to Facebook. In terms of privacy, the participants had 1) concerns about selling of their personal data by digital platforms, 2) a lack of knowledge about the business model of digital platforms, 3) a sense of insignificance, 4) confidence in their mitigation measures and 5) concerns about their financial data being leaked.

The next section deals with the privacy red lines that participants had to the extent that they can share data with digital platforms. This section will also explore the kinds of data participants were comfortable with being shared with a third party.

5.6 Privacy Red Lines

As described in previous three sections, participants interact with digital platforms on a regular basis, but believe that they don't have the necessary tools to engage in actively managing their privacy with regards to the data collected by these platforms. Managing the visibility of their social media posts was the key privacy management practice that was observed across multiple interviews. Hundreds of apps on smartphones made managing privacy a time-consuming task and there was a very low threat perception amongst the participants to allocate the required time and effort. Despite such apathy towards actively managing their data privacy, there were certain data types that the participants were very vigilant about. In this section, such privacy red lines are documented and discussed.

Privacy red lines were the type of data the participants would not share with the digital platforms or will leave the platforms if it was shared with a third party. Sensitive data types varied across participants, bank account or credit card details came across as the most sensitive data and should be protected. G6P4 lists her privacy red lines: “Name and Mobile number is fine, but bank details are not OK. If they ask for bank details its bye bye”. G6P4 would filter out apps or sites that ask for bank details and it was a hard red line that she will not cross. Throughout the discussion, she was suspicious of new apps and if any of these apps tried to cross a privacy red line, she would not install it.

This is similar to the approach that participants referred to in privacy concerns, where they will thoroughly research an app before downloading it to their smartphone. This G6P4 did not consider her phone number to be sensitive data that needs to be protected. Earlier in the discussion, she had mentioned irritating follow up calls from tele marketers after the expiry of a free trial. Other participants like DI3 M27 had a mix of fields like, “photos, your personal photos, contacts, phone contacts, mails, from that your search history.” DI3 M27 had a much more detailed list of data that was considered sensitive, and it included search history. The main reason was her observation that if she makes repeated searches about airfares on a certain route, they tend to go up:

> Financial information. Financial data is secured. If I give access to my email to certain apps, so when there is email statement of my credit card, Google maybe accessing that also. And they can have a, they can access my pattern of spending. So that is a very confidential thing, and it is like I am making my expenditure habits public. I am giving it out to other and I am not comfortable with it. (DI1 M25)

DI1 M25 had outlined financial data and bank/credit card statements as the most important data that needed protection. He had also identified the possible route of leakage via email statements and wanted it secured. The participants felt that no application should ask for financial information or bank and credit card details as that constitutes the most sensitive data. They understood that banking applications, payment applications, e-commerce sites or any other site or application where an online payment is a requirement, will ask for financial details. The second most important data type was photographs, especially for female participants. This came across all focus groups and depth interviews. Smart speakers listening to conversations was a privacy red line that spooked some of the participants:

> When I am talking about tripping somewhere and I want to go on a trip all of a sudden, I start seeing ads That's why on principal I did not buy Alexa I bought a Bluetooth speaker I have turned off Google Assistant for me. (DI2 F23)

DI2 F23 added smart speakers to the list of privacy red lines as connected hardware that is listening all the time seems an added layer of privacy worry. Her main concern was customised advertising based on the listening ability of a smart speaker. The issue of audio snooping and locational surveillance will be discussed in detail in section 6.4.

Another privacy red line was around geo-location data collected by smartphones and shared with various digital platforms. There were mixed feelings about location data; in focus group 4, of males aged 25 to 34, it was felt that location data if used for providing services is acceptable and some of the participants understood that advertising gets better with location data:

> I agree that with location they give more information. Like there is zoning in Newspaper Advertising and you can advertise in select zones only if you do not wish to advertise in the entire state or region. So same is with Facebook, if they track my location information, they can use it to target ads from specific retailers in specific locations like South Delhi. (G4P1)

G4P1 understood how Facebook was using his location to deliver relevant ads. He was also able to relate it to how newspaper distribution works and was aware of how digital local advertising was working. In previous discussions on advertising, this focus group agreed that no one likes ads but also made an exception to relevant ads.

There were some strong opinions about digital platforms tracking their location secretly and, in that case, it became a red line that should not be crossed. G4P4 drew a red line at secret tracking: “Secretly tracking someone I have issues with that. Location is something where I feel a threat to my privacy”. Other than G4P4, location data was not a privacy red line in focus group 4 but was seen as something useful that can be used to deliver relevant information.

Participants in focus group 2 worried about the public availability of their photos. Some of the participants had taken measures to protect their pictures by searching their name on Google from time to time, making their Facebook photos private and some of them removing everything from Facebook. They found Snapchat, with features like self-deletion of photos, very comforting. DI3 M27 mentioned photos as the first thing that he was uncomfortable sharing publicly.

> There was a workshop in a school where in I got aware of the fact that that when you click a photograph the location is automatically attached to it if that option is on the mobile device and be uploaded the same photo somewhere anybody can track that. so, then I changed my mobile setting so that nobody can have my geo-location when I am clicking a photograph or uploading it somewhere. (G2P3)

G2P3 had developed a toolkit where she would disable geo-tagging, turn off comments and the ability of others to download her photos. In her case, the sensitivity was about her photos as well as the fear that someone can track her movements if her photos are geo-tagged. G3P3 was also sensitive about photos: “Our photographs that we do not want to share, now there is an option to selectively share with certain people only”. Selectively sharing photos was one of the main tactics that was deployed to mitigate the fear of misuse of the photograph. This also gave the respondent a sense of control over her digital privacy. In group 3, there were other participants who had concerns about the misuse of their photos, but generally it was felt that the photos were safe on Facebook.

Facebook was seen as a **digital attic** where memories could be stored and retrieved any time. Some participants in group 3 talked about frequently changing phones and this digital attic ensured that the photos are not lost every time a mobile device was changed.

DI5 F32 also discussed her anxiety about the misuse of photos and other personal information. “If somebody is posting pictures and the other person can download the actual pictures without taking permission up that person like pics or number or contact information it's not right.” DI5 F32 believed the data is not safe with any digital platform as she felt that via advertising technology a lot of personal data could be unearthed by marketers. She is a practicing marketing professional living in Mumbai. She does not have either the Facebook app or the Google search application on her phone. Her mistrust arose from a position of being informed about the potential of advertising technology to gather details about individuals. In group 6 (females aged 25 to 34), participants talked about restricting the access to their photos to friends of friends and not everyone and G6P2 was very specific about not giving permissions to access her photos to new apps. This is another example of participants applying a trust threshold to apps and only those apps that pass this threshold get access to their photos.

This section lists the privacy red lines that the participants do not cross easily. These redlines were enforced for new apps and were also used to filter out any new requests. These redlines were 1) New apps seeking financial information like bank accounts or credit card details, 2) giving access to photos, 3) giving access to contact lists, 4) secretly tracking of location and home address, 5) smart speakers always in listening mode and 6) giving access to emails. Apps that had passed the trust threshold were less likely to be filtered over these concerns.

The main controls that were applied to mitigate privacy concerns were 1) disabling geo-tagging of photos by cameras and friends, 2) restricting access of their photos to a select set of friends on Facebook, 3) disabling comments and options to download photos, and 4) employing a trust threshold to apps that have access to their photos. Some of the participants were happy with the

measures for protecting their photos and were pleased with using Facebook as a digital attic where all their photos were stored, while others had removed all their photos from Facebook. Most of the references to photo sensitivity came from females and this sensitivity was expressed across cities and age groups.

Davies (1997) described neutral language like loyalty cards that were used in the past to encourage people to sign up and share their transactional data with retailers. Turow et al (2015) investigated the practices of combining such transactional databases with online databases to create a complete profile of a shopper and uncovered a network of relationships between online and offline data brokers. These relationships between indexed computerised databases brings forth an unsurmountable challenge for the consumer where they must manage their privacy in an environment that is largely invisible. An invisible predictive network serving personalised advertising is a challenge to privacy management and gives rise to folk theories about the source of private information.

The next section investigates the privacy policies and discussions around consumer data privacy within marketing departments of advertisers, media agencies and publisher platforms. This is based on interviews with Industry Experts that were recruited because of their 10 plus years of work in the Media and Communication industry in India.

5.7 Privacy Practices in Indian Organizations

The findings in this section relate to RQ2 and Industry experts were asked about the process of creating custom audiences and the state of user privacy discussions in the digital marketing process. They were also asked about the changes that have happened over the years as they have shifted from traditional marketing channels like Television and Print to digital platforms like Google and Facebook. Participants from media agencies, marketers and media platforms were asked about the practice of creating digital media campaigns and the discussions around user data privacy during the planning and execution process.

Intelligent Tracking Prevention (ITP) is an apple initiative that restricts use of third-party cookies to track internet users across the internet. It has been becoming increasingly restrictive in the use of third-party cookies since the launch of this program in 2017. When asked about it impact on the media planning process in India, Industry Expert 3 who works in a media agency in Delhi said, “From an agency perspective ITP is not a big factor. This might be due to the fact that in India as market, iOS devices contribute to less than 5% so from an advertisers POV its not relevant”. Industry Expert 3 explained that there were no discussions happening in his agency about ITP that was a prominent topic in the Western countries with high penetration of iOS.

Intelligent Tracking Prevention (ITP) that was high on the agenda in the Western world was not of high importance as India. The main reason is the India is an Android-dominated market and most of the strides in tracking prevention have been initiated by Apple in its Safari browser. This is an important descriptor of a privacy divide that has started building between geographies depending on the predominant mobile operating system. Similar European General Data Protection Regulation (GDPR) also did not feature in the data decisions of Indian marketers. Industry Expert 3 describes global clients as being more aware about data privacy issues as compared to clients with clients that were local.

> Considering that most of GDPR Fiasco was limited to EU some of the global clients yes they are actively talking about it, seeing how it's affecting their business in India but there are some clients who are way behind on the Curve and for them privacy is a very new concept it only gets tackled when they reach that place for example any client of mine right now a privacy concerns will be very very last day that they would be considering. (Industry Expert 3)

Industry Expert 3 elaborated that the differences between attitudes towards data privacy can be attributed to the ownership structure of the marketer. In his opinion, multinational organizations were slightly more inclined to have discussions about user privacy as compared to Indian organizations. In both instances, the discussions are not accorded great importance for media planning and execution of digital campaigns. According to Industry Expert 3, General Data Protection Regulation (GDPR) was a very European phenomenon, and the Indian industry was largely unaffected by those discussions. Advertisers were not actively seeking to understand the privacy issues related to the source of rich consumer data that is used for media planning:

> Discussions have just started it’s too early and it's not a big concern right now it is not affecting things as of right now, but the larger companies are being conscious before any kind of regulations come in, but other than that it is not a major concern. (Industry Expert 5)

Industry Expert 5 is a CXO for a digital media agency in Mumbai and he confirms the observations of Industry Expert 3 that consumer data privacy is not a topic that is actively discussed during the media planning or execution process. Industry Expert 6 who works in a broadcasting organization in Mumbai described the consumer data privacy discussions in the organization as: “Never ever do we discuss in the organization. The only thing that comes up is the food pledge but that has nothing to do with data privacy”. Industry Expert 6 works for a television channel that has a nascent digital business but privacy discussions within the organisation are rare.

> And user data all these were taken very seriously as employees we were pretty much sign undertakings once a quarter once at 6-months about user privacy. I'll probably stick my neck out and say 95% of the world out there will not care about receipt of consumer data or anything of the sort. (Industry Expert 7)

Industry Expert 7 who works in Mumbai, has experience across television and digital platforms. He explained the difference between global tech giants where there is continuous discussion and training about consumer data privacy and the rest of the industry where there is no discussion at all. This view is corroborated by Industry Experts 3, 5 and 6.

Nguyen and Solomon (2018) identified seven modes of passive data collection 1) cookies, 2) web beacons/pixel tags, 3) device information and tracking, 4) fingerprinting, 5) facial recognition, 6) payment cards and loyalty cards and 7) data collection from third parties. Industry experts that were interviewed for my research were able to outline the use of 1) cookies, 2) mobile advertising IDs like Identifier for Advertisers (IDFA) for iPhones and Google Ad ID for android devices, 3) third-party data management platforms (DMPs), 4) website analytics, 5) advertisers' customer relationship management (CRM) systems, 6) third-party lead aggregators, 7) mobile services operators, 8) online sales or e-commerce, 9) second party affinity data from Google, Facebook and other digital platforms and 10) loyalty cards. The use of credit cards was not discussed in an Indian context and the focus was on the loyalty cards of big retailers and multinational corporations like Starbucks; however, data aggregated by telecom operators came up as big source for creating audience profiles.

Industry Expert 1 who is a CMO in an Insurance organization describes data privacy policies: "With the sales team we have a very robust compliance received the reviews before we reach out to a customer base that was targeted a few months back. There are guiding principles about that and we do follow that". Industry Experts 1 and 7 talked about strong practices around consumer data privacy and explained that overall, in the marketing fraternity such discussions are not the norm. Industry expert 1 is from the insurance industry that is one of the most regulated industries in India and is used to handling sensitive consumer data and the same practices have carried over to digital marketing. However, Industry Expert 1 was able to comment on data that was available from telecom providers for profiling as well as targeting of potential customers for insurance products:

> They do integrate with Google Analytics and lot of third-party data authentic third-party legally available data in terms of persona building so lot of telecom companies do it because they know everything about the customer. He will not give me the PII data I go to a Telecom operator say Orange he says I have a database off 100,00,0000 customers

> now he exactly knows what is the consumers monthly billing, how much time he spends on the internet. Now you also have tools where you can see what kind of SMS's he receives. Technically if you ask me, I can even know what is your salary because the minute your salary is credited you get an SMS I know what you are getting., when you are paying your credit card bills so I also know your out flows. (Industry Expert 1)

A telecom operator with millions of customers that can be tracked across time and geography combined with accurate financial transaction data trails can operate one of the most accurate profiling businesses outside the realms of technology platforms. According to Industry Expert 1 these sorting, profiling, and targeting capabilities were available to marketers in India and used by marketers like him.

Similarly, Industry Expert 7 was able to provide a glimpse of profiling and targeting information available from a global digital platform:

> At a corporate level all the data is shared with the advertisers is at an aggregated level. Even at an aggregated level it's very specific to a platform. If I have to speak for Platform X it will be about a job function, it could be industry, it could be seniority and will be very relevant to the specific advertiser. Then you have a platform like Google and DoubleClick where you can go deeper into specific aspects. What is the kind of demographic in terms of what is the kind of gender, geography etc. and of course there is Facebook that talks about your interests and likes and that aspect. I think data is there and advertiser can literally build a persona of anyone out there, with the kind of aggregate data that it gets. (Industry Expert 7)

Industry Expert 1 provided a detailed example of how a telecom operator can create audience profiles that are suitable for the insurance industry and the process of targeting these profiles using advertising technology, where data from multiple sources is aggregated for "persona building". Whitaker (2000) and Solove (2004) discuss the process of creating data doubles of individuals through linked online databases that were made possible with the advent of advertising technology. Industry Expert 7 described the creation of data doubles by aggregating information from multiple platforms like Google, Facebook, and LinkedIn. Castells (2007) described the value of information and its impact on how it shapes the world around us. All the Industry Experts were able to provide details about how they create insights from information that is derive from the data that is collected. Thrift (2005) incorporated the issue of location into privacy where devices are aware of the location of the consumer and the implications for consumer privacy.

Mayer-Schönberger and Cukier (2013) emphasised the importance of big data comprising of information. The scholarship converges on the centrality of information extracted from datafied individuals as the fuel that powers information technology. Industry Experts had shared detailed processes and understanding of the role of information for all sections of the digital ecosystem from advertisers, media agencies and digital media platforms.

Mayer Shoenberger and Cukier (2013) also provided an alternative framework for safeguarding the privacy of users by suggesting that regulators make digital platforms responsible for the secondary usage of consumer data. According to Industry Expert 7, multinational digital platforms took the responsibility of securing user data very seriously. Other industry experts described little to no discussion about privacy in the entire process of creating digital media campaigns. Industry experts knew about GDPR, but it was viewed as a European issue that was starting to impact the Indian media and advertising industry

Turow (2011), Young (2018) and Kaplan (2020) discuss the creation of postdemographic audiences where advertisers are given access to richer profiles of platforms users to create an accurate profile of users. Industry experts outlined detailed practices of how postdemographic audiences are created. The complex layer of advertising technology and relational databases were largely invisible to Everyday Users, and they were unable to fathom the accuracy of the advertising in targeting them with highly personalised advertising. Wolff (2017) described the increased value assigned to audience attention to advertising and the role of digital media buyers came up because of this change and the underlying linkage of the behavioural and transactional data that is possible only for digital platforms. Young (2018) outlined the different data trails that constitute the audience commodity that is traded by the media buyers described by Turow (2011).

Carmi (2020) described the role of advertising cookies and related aspects of advertising technology that transports and links information about cohorts that is useful for aggregating audiences across the internet into advertising-friendly groupings. As seen above these data groupings are also called data dossiers. According to Cheney-Lippold (2011) data dossiers need to be sorted and individuals classified, so algorithms play an important part in this classification. According to Carah and Brodmerkel (2020) the objective of this classification is to present the advertisers with the opportunity to integrate into the digital lives of prospective consumers. Industry experts from digital media agencies and advertisers were able to identify opportunities offered with data rich profiles and the data sources. Some participants were not able to describe these processes and had to interpret the source of personalised advertising through vernacular theories.

The process of advertisers integrating into the digital lives of consumers raises questions about consumer data privacy and Westin (2000) brought up the privacy divide that exists across socio-economic class. Shoenberger (2017) discusses the privacy paradox where there are voices in the media and surveys that talk about the importance of privacy but hardly any purchase or usage decisions get made based on privacy. Some of the Industry Experts were able to elaborate on the impact of privacy initiatives in the EU and by Apple, but they felt that without a privacy law and a very low penetration of Apple devices the issue of consumer data privacy was dependent on the advertiser.

Big data collection is not a new phenomenon and financial data gathering for the purpose of providing credit scores is as old as the banking system itself. Gandy (1996, p. 139) listed several sources for personal information and computerised databases that existed prior to the mass adoption of mobile internet and smartphones. The advent of digital technology has enabled the creation of relational databases where financial data, purchase data, location data can be combined with digital consumption data to create powerful signals for advertisers to segment and target audiences, with little or no control exercised by the users.

In the literature Solove (2004, p. 23) called the profiles created by processing digital data trails “digital dossiers”, and listed multitudes of sources where the data was stored and could be retrieved to target visitors to websites. In my research Industry Experts from marketing functions in insurance, the beverages industry and media agencies were able to articulate industry-wide practice for creating audiences that were suitable to be targeted. Participants across multiple focus group discussions exhibited awareness of their data being collected and used for advertising. Participants were not able to fathom the amount of data that is collected and sorted.

Solove illustrated the impact of technology on privacy through Focus USA’s 100 targeted mailing lists encompassing information about people’s credit history, affluence, ethnicity, life stage and mobility. In my research Industry Experts were able to identify their target audiences based on their career stages, real time in-market behaviours across geographies and product categories, food and beverage habits and routines, and purchase journeys for cars and insurance. This real time profiling and targeting exhibits the key change between 2004 and 2019.

Raley (2013) describes three dimensions of dataveillance as descriptive, predictive, and prescriptive. In my research, Industry Experts were able to pinpoint the importance of identifying their consumers across online and offline avatars and the role of first-party, second-party and third-party databases to create a rich profile of digital consumers. Examples that were provided by the Industry Experts were profiling of consumers using their eating out habits, their search history for insurance products, aggregated profiles of consumers of telecom operators, activities

on digital platforms, consumer affinities built on their online behavioural data and geo-location. Raley (2013, p. 126) envisaged a “superpanopticon” system with the ability to coerce the desired behaviour out of data subjects either by the state or corporates. Some participants in my research did not feel that they were exploited by the digital platforms and took a more affirmative approach towards the benefits provided by the mobile internet technology.

Some Industry Experts were able to confirm that deliberations on user data privacy were not institutionalised. Multinational advertisers that have an exposure to Europe and GDPR were aware of the developments in the field of consumer data privacy and were discussing the impact of advertising on user privacy. However, as most of the initiatives were driven by Apple, the impact on the advertising technology in India was minimal.

McStay (2017) described the process of real time bidding that makes it possible for advertisers to participate in a real time auction for an opportunity to serve an advertising impression to their target audience. This target audience is generated based on transactions, online behaviour and the predictive power of ad serving algorithms is enhanced by a big online presence of the individual. Tsesis (2018) outlined the vast sources of databases that can be accessed and are used to profile an individual across multiple dimensions depending on the interests of advertisers. Comparing the findings of this research with McStay (2017) and Tsesis (2018), it is evident that in terms of the advertising practices of digital media, Indian media agencies were following the global processes with regards to the utilization of advertising technology. The major difference was the awareness of issues around user privacy and the incorporation of privacy into marketing functions of most of the organization.

5.8 Conclusion

Some of the Everyday Users were unaware about the possibility of third parties aggregating their data and the use of ad tech to create a detailed profile from disparate sources. Financial data like bank account details, credit cards details, phone contacts, ages, photos, phone numbers, home addresses, leisure activities, and search and watch histories were found in multiple mentions as sensitive data. Female participants across cities and age groups were very sensitive about public access to their photos. Some of the mitigating measures that were undertaken included restricting the visibility of social media activities and employing a trust threshold to all new app downloads.

There was a sense of resignation that the participants were too insignificant to be able to negotiate with the digital platforms if they wanted to use the apps. This reflects issues that Westin (2000) also raised; consumer control on their data and the ease of exercising this choice.

Westin (2000) also brings up the aspect of consumer privacy and surveillance by describing the delicate balance about the quantum of data required to provide a given service to the users versus the amount of data required to provide the service. Nguyen and Solomon (2018) found out that consumers of digital platforms lacked the skills to engage with privacy policies and control their visibility. They also discovered a lack of research on the impact of advertising technology on consumers. My research fills this existent gap in essential research into the understanding of data practices, with a focus on practices of creating personalised advertising by combining digital data trails across various sources.

There was discomfort amongst participants about lack of control over data privacy. This did not result in any changes to their everyday practice. This lack of action on privacy concerns was uncovered by Shklovski et al (2014) and labelled as privacy paradox by scholars such as Shoenberger (2017) and Solove (2004). This privacy paradox was also encountered in my research.

My project uncovered the method of self-audit of preferences and the ease with which participants were able to discover how privacy control can be developed into an easy-to-use toolkit. This toolkit can be delivered through regulatory, educational, or citizen-led initiatives to digital consumers for enhanced data literacy.

None of the participants were able to recall explicit endogenous communication from platforms about changes in specific privacy controls or guidance about where to find such controls. Participants presented three main reasons for not paying attention to the data usage policies, 1) obfuscation of important information in EULA, 2) TINT and 3) TINO. Participants outlined their anxieties about the possibility of digital platforms selling their data to third parties.

This chapter examined the level of discussions that were happening around consumer data privacy in Indian companies and focussed on the depth and breadth of information available for marketers in India to construct and reach their target audience. According to Industry Experts, the level of discussions across advertisers and agencies were miniscule, if at all. The only exception were industries that were regulated by law or multinational technology companies that are following global protocols. Raley (2013) describes various practices of the digital advertising systems to create audience profiles before the fact and Industry Experts were able to describe how first-, second- and third-party data is used to create profiles of digital audiences that are used for customised advertising. Participants expressed their sense of helplessness expressed as TINO, for all or nothing approach of apps where they must share all the data that is asked for in exchange for access. Industry Experts described how behavioural and contextual targeting could interpret important aspects of a user's life stage and serve customised advertising that is

extremely intimate to their personal situation. A lack of understanding about the underlying layer of advertising technology made some users uncomfortable when encountered with personalised advertising messages.

Next chapter explores whether users experience advertising technology as surveillance, an enabler to access critical services or both. Some key concepts such as location and audio snooping are described, and an effort is made to understand if the participants feel exploited by being sorted and targeted by advertisers on digital platforms.

Chapter 6 Surveillance

6.1 Introduction

Digital trail data collected from the users of digital platforms enables advertising technology to create interest and affinity driven audience profiles. This exhaust is also described by Hepp et al (2018) as “digital traces” left behind which are the online activities of individuals and their social group that contribute to the aggregated big data held by technology platforms. These digital traces enable dataveillance and advertising surveillance. These profiles are used to track and target users across the internet for the purpose of delivering advertising that is intimate to the circumstances of the users. To understand the customised messages that are delivered to the users of digital platforms, it is essential to explore the marketing processes that are used to create these profiles. This surveillance is surfaced to the users through advertising that is intimately aware of their current interests and affinities. Advertising that is aware and accurate drives the users to an uneasy interaction with a complicated information environment.

Elmer (2003) described the use of information and advertising technology deployed to track the journey of an individual user across the internet. According to Elmer (2003), the main tool for tracking users across the internet is the cookie and since its humble beginnings in 1994 it has become the mainstay of online advertising surveillance. Wall (2006) pointed out the unregulated use of cookies that existed and how a small piece of code that was designed to provide better customer experience to repeat visitors of a single website grew into the mainstay of advertising technology. McStay (2009) describes the practice of gathering big data about millions of consumers as dataveillance which is the fuel that powers the advertising funded digital services.

This chapter explores dataveillance through the architecture of data flow between various entities in the advertising technology ecosystem. First half of this chapter investigates how the participants receive customised advertising. Industry Experts describe the practices used to create qualified target audiences for their marketing campaigns. They outline the process of audience profiling, audience targeting and the capabilities of ad tech in serving personalised advertising. Everyday Users in my research describe factors that shape their attitude towards customised advertising and the role of advertising in delivering useful information. Second half focuses on the perspectives of participants about advertising surveillance and highlight their experiences with this complex ecosystem. Towards the end of the chapter various folk theories about data collection are examined.

6.2 User Profiling and Customised Advertising

User data is critical for media planning and for digital platforms that mediate the conversation between marketers and their end users. Industry Expert 1 is a chief marketing officer (CMO) with experience across multiple industries. “I would say there is an overflow of data in terms of especially when you are doing planning on the online advertising because over the years data has become very rich and platforms are much richer.”

Industry Expert 1 was very satisfied with the amount and depth of data that was available to him to plan digital media campaigns. He felt that the quality and availability of second-party data had improved over the years. He went on to describe an elaborate system that brought customised digital advertising to life. The most important manifestation of this process is the audience profile that is created for every campaign using multiple sources of data.

Industry Expert 3 worked for one of the big media agencies that specializes in digital media planning and buying. Industry Expert 3 lists three main sources of data that are used for the purpose of digital media planning. “First-party data from me as an advertiser what I have, to secondary data that Google and Facebook give to me, plus third-party data providers for example DMPs”. Data management platforms (DMPs) collect data from multiple platforms and enables advertisers and media agencies to create detailed profiles of audiences. Elaborating on the role of third-party data providers, Industry Expert 3 explained the process of data collection and attributes that are used for campaign planning:

> Apart from age and gender, their habits, the role of data partners is to place SDKs inside apps. The data partner would have details like location, gender, age, consumption pattern in terms of what kind of apps that a particular device ID has on its phone. What is the kind of usage that a user will be seeing on these apps. The location combined with app usage gives us an ample consumer profile to go after consumers. (Industry Expert 3)

SDKs are software development kits that are integrated with the applications that people download on their mobile phones for a particular purpose. SDKs collect usage data from mobile phones along with an identifier that is used to aggregate this data with first- and second-party data that is available with the media agency to get a complete picture of a consumer. First- and second-party data provide insights about consumer activities within the walled gardens (Imagine Facebook, Google Search or TikTok), whereas DMP provides data about the activities across the internet.

The first step in the process of creating a target audience for a brand campaign is the integration of these data sources into a targeting mechanism and selecting the best platforms to deliver

advertising to the desired cohorts. The whole process at the end of the media agencies is geared to answer business questions of their clients. Industry Expert 3 goes on to explain the process involving owners and the role of a media agency in planning and executing the communication plan. “It is kind of the question that the client raises. It's the kind of marketing and the kind of objectives that the client would want to work around with. This is what the agency needs to tackle.” (Industry Expert 3)

In my research, industry experts made numerous references to the importance of internal practices to manage consumer data (first-party data) and use it to map the product purchase journey. Industry Expert 1 highlighted multiple information sources that are explored by consumers on their path to purchase. Industry Expert 1 also outlined an extensive array of industry practices that are employed to map sources of data to different stages of the purchase journey:

> Let's see that from online where does the customer get the information? The product page of the organisation, there are a lot of influencers, a lot of bloggers who influence certain products you have lots of aggregator websites for example a CarDekho and all who give reviews and you can do comparisons across products, their people to put in their Leeds so if I'm interested in the car I will put in my lead. So apart from the organisations pages you have a lot of these pages where the first link is ok you can ask for a test drive. Now the entire journey is mapped. Whether he came from a third-party website or whether came from a company website and he has said OK I'm interested in a test drive and there it goes on the offline mode. So, there you can actually track how many people actually came and I think most of the people today at least the Urban India book a test drive before walking into the showroom. (Industry Expert 1)

The user product purchase journey can now be tracked from the time they first interact with the product like the website of an organization, video site, social media site, search app and lead aggregators like CarDekho.com. The user is linked through all these touchpoints enabling advertisers to map multiple journeys and there is a competition among various advertisers to nudge the user towards their brand.

Combining online and offline data about individual consumers enables a car brand to track the journey of a customer from information search stage to a visit to a physical car dealership to take a test drive. This process provides the marketer with a conversion funnel starting from a website visit and ending with a test drive. Various drop off points in the funnel provide the marketer with an opportunity to communicate with the consumer and try and get them to engage with the

brand. Integrating data from online and offline sources provides highly valued conversion metrics and the role of advertising in guiding consumers through the purchase journey.

Depending on the advertising budgets of competing marketers these conversion metrics can drive aggressive retargeting. Participants in previous sections have described instances where the ads continue to follow them across the internet even after they have made a purchase. Across multiple interviews Everyday Users in my research have recounted how a search on one platform triggered advertising of a totally different platform. These experiences surface the interconnected nature of advertising surveillance. In some instances, the users are able to link an accurate advertising delivery to an activity on the internet, while in other instances they are perplexed at the accuracy of advertising leading to formation of folk theories.

The insights generated about select users that move from being a casual website visitor to someone seeking a test drive can be used to target a much larger pool by creating lookalike audiences on leading digital platforms:

> You can always keep enhancing your media plan as compared to traditional Media. Data is there, I would say the only challenge when it comes is when you look in inwards, because a lot of the times as an organisation we may not have that much of data about our own customers but when I do customer persona building and then I have to look for the lookalike then platforms and data is available. (Industry Expert 1)

Industry Expert 1 found the immediacy of course correction on digital platforms a big improvement over traditional media. He also makes an important observation about organizations creating internal systems to gather and store first-party data. This data can be used to find similar audiences from a much larger pool using advertising technology. According to Industry Expert 1, there is an abundance of available data, but there is a requirement for carefully designed data practices for optimal utilization of second- and third-party data to enrich media targeting.

Advertising technology relies heavily on various identifiers to track individuals as they move between the virtual and the physical world and it is also enriched by meta data that surrounds the data thus generated. Lyon (2014) defines metadata as data surrounding the data that is being gathered. Mayernik and Acker (2018) described the opaque nature of metadata and Everyday Users in my research had no knowledge of metadata that was generated by their usage of digital platforms. Glass and Callahan (2014) named the combination of postdemographic data and metadata as the “data muscle” of tech companies that own the digital platforms and flex this muscle through deployment of product features and personalised recommendations. Flyverbom

(2019) shines a light on the “big data” that gets accumulated due to the linking of user behaviour across online and offline activities. McStay (2018) describes the use of behavioural and transactional data for advertising and changing behaviour of product consumers through advertising.

Based on discussions so far, the process of advertising surveillance begins with matching audiences of digital platforms with the target audience required by the advertisers. Advertisers share details about their desired audience and the media agency combines data from multiple sources to arrive at an optimum mix of digital platforms to target them. Apart from media agencies that manage the process of targeting potential customers online, marketers also engage third-party lead generators who perform an important function enabling comparisons for industries like cars and insurance. Industry Expert 1 described the process of lead generation and remarketing in the sales of insurance products:

> PolicyBazaar an online platform which is supported by a call centre so in case I have a query you know ok I like this product and have certain queries you can speak to the call centre you can chat on the WhatsApp they have integrated all these communication channels and created a single API. If you drop off during your purchase journey you will get a call, the link on WhatsApp so we can just click on the link and it will take you back to the page where you dropped off and you have a bot which will say what was the issue, why did you drop off, what was the information you wanted. If you're not happy with the bot you can say call me so within 2 minutes you get a call. (Industry Expert 1)

The process described by Industry Expert 1 has a high level of responsiveness and feedback built into it. There is an audit trail to understand the drop off points and multiple options for the customer to re-engage with the process. The phone number of the customer becomes the key that identifies the customer, from information request to purchase and provides an avenue to reengaging with the customers who drop off along the path to purchase. Some of Everyday Users did not identify their phone number as sensitive data that they would not wish to share with various digital platforms or for getting a free trial. As stated above, a phone number is generally used to track users across the internet.

This process of converting consumer data into targeting signals for advertisers exhibits the level of visibility that marketers have of users across digital platforms. This process also shows the integration of online advertising to offline marketing apparatuses like call centres and third-party lead aggregators. Industry Expert 5 is a CXO for a digital media agency and he describes a process that is similar to the one outlined by Industry Experts 1 and 3:

> Data can be collected by multiple entities, can the data come together? Yes. Advertisers collecting first-party data and agencies collecting third-party data, for example Interest level data. Both can be matched together for some advanced targeting. For the web world it can be a cookie sync process to sync that data and bring it together. In the app world GAID could help you sync up that data. So, for example this insurance guy is getting data for demographics, that can be paired for identifying that user, what that user is doing for example through the agency data. You can identify certain users that are about to travel. While this person has taken motor insurance from you, car insurance and you know that the guy is about to travel abroad, now you can cross sell your international travel policy to that person. (Industry Expert 5)

Industry Expert 5 gave an example of how an insurance company could employ first-party data to identify an immediate need and reach the prospective customer with topical advertising. This data integration creates a technology platform that can help advertisers predict future behaviour of customers with an affinity for their product category. On the other hand, the user might understand how the insurance marketer know if they are about to travel. These are the unexplained deployment of advertising technology that create disquiet among users.

Multiple scholars (Gandy 1989, Dandeker 1990, Bogart 1996, Gilliom 2001, Campbell and Carlson 2002, Elmer 2003, Mayer 2009, Saulles and Horner 2011, McStay (2017), Szulc 2018, Carmi 2020) have compared the data aggregation and data activation practices of organizations in retail, banking and finance, travel, insurance, welfare, taxation and big technology to panoptic surveillance. These scholars have used Foucault's analysis of Bentham's panopticon to investigate mass surveillance using computerised databases. Advancements in ad tech were a direct result of improvements in the quantity and quality of online behavioural data and identity solutions described by scholars (McStay 2018). Industry Experts in my research were able to describe how a combination of behavioural and transactional data has created advertising systems that deploy predictive analytics to target audiences based on intent.

There are three main differences between the use of Foucault's analysis to classify digital advertising as a tool of manipulation, and deployment of personalised advertising for driving the business results of advertisers. The first difference is the inability of the advertisers to inflict punishment for non-compliance. The second difference is the ubiquitously invisible nature of advertising technology, whereas a panopticon derives compliance through the perception of being watched. The third difference is the centralised nature of observation in a panopticon as compared to the decentralised nature of advertising surveillance with multiple brands in the same category trying to influence a target audience.

An optimal utilization of consumer data requires an ability to track transactional and behavioural data and assign it to individuals. GAID is Google Advertising ID and its purpose is to help advertisers track consumers over time and across different digital platforms:

> It could be done on cookie sync, it could be done on email matching, and it could be done on phone number matching, there are multiple options available. There is also GAID Google ID so all those can be linked up. Typically, someone would use the ready DMP platform, and you start collecting data there, so an advertiser could have a licensed DMP at their end so can an agency have it, there are also third parties who can help sync up the data. To bring it together and ultimately use it for targeting. (Industry Expert 5)

Industry Expert 5 emphasises the importance of identifying the consumers and deduplicating them into a Data Management Platform (DMP). This means advertisers create their own infrastructure to exploit the data that is generated with every interaction with a digital asset. Cookies, phone numbers and GAIDs are used to identify a consumer across time and devices. This integration enhances the ability of advertisers and their media agencies to deliver highly customised advertising to the most desirable customers in their product category. Such behavioural and contextual targeting of audiences is possible if there is close alignment between advertisers and media agencies.

Similar to the processes and identifiers outlined by Industry Experts, McStay (2017) listed identifiers that are used to link consumers across websites and devices. Industry participants in my research shared a similar list of identifiers like Apple's Identifier for Advertisers (IDFA), Google ID, Phone Numbers, email addresses and IP addresses that can be combined to track user behaviour across the internet and across multiple devices that a person may use.

Industry Expert 2 works in Delhi and is the CMO of a food and beverages company. He talks about the immediacy with which data can be acted upon in digital planning and the depth of data provided by platforms:

> Television is not interactive digital is, so you can have chats with your consumers then and there that's important you can gratify them immediately so there is a bigger connection to digital which is not possible to television Media and the other kinds of data that you'll probably get where they are from location perspective what are they viewing the restaurant they are in because of all these apps and all these locations. I don't know whether they will be called intrusions the platform gets to know where the consumer is. (Industry Expert 2)

His line of business is very dependent on the location of the customer as his beverage sales are dependent on bars and restaurants serving them. An integrated advertising system that has spatial, demographic, and behavioural data helps Industry Expert 2 in executing his marketing plan with fast turnaround times as compared to traditional media like television. Knowing a consumer up close and personal is an attribute of digital media that is unmatched by any other medium. Second-party data coming in from top digital media platforms is the most important factor in choosing the right vehicle for his advertising message. This spatially aware advertising system is also referenced by G5P4 in the previous section.

The computerization of transactional and behavioural data and its storage in easily retrievable databases has created some of the major challenges for consumer data privacy and control over their own data. Bogard (1996) outlines the importance of profiling as the first step towards creating a surveillance system and relational databases make it possible to create detailed profiles of individuals and aggregate them into cohorts. Whitaker (2000) described that the information thus generated is the product that digital platforms sell to advertisers as postdemographic audiences.

Industry Expert 1 reflected on the changing trend on data availability and other experts shared similar experiences where the amount of consumer data available for them to create detailed targeting profiles was increasing exponentially over the years. Inputs from Industry Experts 1, 2, 3 and 5 provide a comprehensive overview of different sources of consumer data that are used for digital media planning. The main reason was the increased adoption of digital platforms strengthening their “data muscle”. They also emphasise the importance of advertisers to invest in developing advertising technology infrastructure that enables the integration of different streams of data that are linked together via multiple identifiers.

These discussions have enabled an understanding of the process where digital trace data from multiple sources is integrated to create a profile of audiences that is used for targeting. Delivering customised advertising is where all the aspects of datafication converge. The next section explores how customised advertising enabled by advertising surveillance is experienced by the users. Experiences of participants while interacting with a complex profiling and targeting system that is highly predictive of their immediate personal preferences will be enumerated.

6.3 Attitudes Towards Advertising

This section documents the attitude of participants towards advertising and the impact of the context on how advertising is perceived. Another focus area of exploration will be attitudes and

the understanding of recommendations and customised advertising. The role of platforms in shaping attitudes and perceptions will also be explored in this section.

Advertising on Instagram came up for praise from across multiple interviews due to its compatibility with the native format:

> I do click on ads. If I am surfing through Instagram and I see an advertisement if I like it a lot and I want to see what is it about. I always open it and most of the time it happens. I have purchased stuff after seeing those ads. I just bought a shoe when looking at the ad only. When I was looking at the site, I was so fascinated I just ordered it. (G2P2)

G2P4 voiced support for personalised recommendations as “Instagram, they know my information, so they show me that particular thing that I like.” G2P4 appreciated that Instagram used the information that it had on her preferences to show her relevant advertising. G2P2 not only liked the advertising that is served on Instagram, but also bought products advertised on Instagram. This speaks of the ability of Instagram to serve highly relevant advertising. Both participants engaged with advertising on Instagram and provided an endorsement for the format and relevance of the ads that are served to them. It was important for the participants that advertising was rendered to be endemic to the content on the platform. Though G2P2 did not directly express the point about the relevance of advertising. However, she detailed how advertising on Instagram led to purchases. The ability of Instagram to find the appropriate context to deliver the message come out strongly.

In the same focus group, G2P3 spoke about relevance on multiple occasions: “You are shopping for a brand and if they're showing similar products, it's useful. So that's fine.” G2P3 also supported personalised advertising “I mostly use Instagram and they show relevant ads when I'm searching for something”. G1P1 found relevant ad personalization beneficial: “when I am in the phase of buying something and I'm looking when I get ads served, for example I'm looking for a Bose speaker and I get ads for JBL speakers I don't mind.”

As seen above some participants across groups (G2P2, G2P3 and G1P2) found advertising to be useful information that enabled them to make a better purchase decision. The participants mentioned Instagram on numerous occasions when talking favourably about a likeable advertising experience. This mention of Instagram as an advertising platform was unprompted and indicated a perceptible impact it had on the participants. DI3 M27 liked advertisements only on Instagram:

> Instagram is good because it's like the ads are more varied over there you can get ads in terms of electronics, apparels, travelling, all kinds of ads are coming over there and I have not seen repeated ads over there. (DI3 M27)

DI3 M27 listed three reasons for a favourable opinion of advertising on Instagram: 1) a wide variety of advertisers, 2) advertising customised to his interests and 3) limited repetition of advertisements. Compared to other platforms like YouTube, a manageable frequency of advertising was a positive for Instagram.

As mentioned above, relevance comes across as very important for multiple participants. DI3 M27 did not prefer advertising on YouTube because it interfered with his primary purpose of video entertainment. “In YouTube it pauses your video what you are seeing and plays the ad.” Participant was frustrated with Facebook as well. “I have stopped using Facebook entirely. It was irritating in terms of ads, more ads were repeated and not targeted”. Apart from repetitive exposure to the same advertisement, exposure to irrelevant advertising was another factor for him leaving the platform. DI3 M27 brought up the expectations that participants have from advertising and how different platforms perform against those expectations. Instagram comes across as the preferred platform for receiving advertising. Advertising on YouTube was disliked as it interrupted the experience of video entertainment, whereas Facebook was disapproved because of the volume of advertising and lack of frequency control.

DI5 F32 Mumbai expressed a positive opinion about search advertising as it aids a task-oriented activity whereas advertising on YouTube or any other video platform was viewed as a hinderance with little or no control over the experience:

> When you are researching or doing something on Google and some ad pops up its fine. If a user closes it should not come again. If at all in YouTube or Netflix or Amazon Prime you are watching something like a movie and unnecessarily an ad comes then that's a problem. (DI5 F32)

Ads interrupting a long form video viewing experience was found to be generally annoying by participants. G5P2 preferred ad-free video entertainment apps like Netflix for this very reason:

> The reason we are on Netflix so much is because it is advertising free site as of now. If I go to Hotstar and I see a 30 second ad it gets a bit annoying. If it is skippable ad, we are OK with it. (G5P2)

If the desired video content is on a platform that is advertising-supported, skippable ad formats were preferred over platforms that do not offer an option to skip the ads. Usually, platforms like Hotstar, that are owned by a broadcaster, follow the television model of a fixed break format with no option to skip the commercials. YouTube does provide an option to skip some of the commercials, but there are short duration advertising formats where the option to skip is not available.

Google Search and Instagram came up as favoured digital applications with regards to advertising customised to the context of surfing. Generally, participants liked the fact that platforms were showing **relevant** ads; the nomenclature used across interviews was "**targeted or customised**". Participants liked ads related to searches that they were performing. The advertising that helped the consumer in making a choice or provided more information around products or brands were appreciated. However, not all ads were appreciated and the platforms where **ads interrupted the primary experience** were disliked. Other ad formats like display and banner **ads that could be ignored** were easier to tolerate. G5P5 made a direct comparison between video and display ads:

> The other forms of ads like banner display etc does not bother us as they can be ignored. The thing that bothers us is the video content that you are watching, and something pops up. That is very annoying. (G5P5)

Participants showed a positive disposition towards digital platforms that do not serve excessive and repeated advertising. They preferred their entertainment to be uninterrupted by advertising and this was not possible when they were consuming videos on **advertising-funded digital platforms (AFDP)**. In this case, the level of annoyance can be mitigated by providing some level of control over the advertising delivery.

Some participants were not happy with advertising interruptions on YouTube, though there were some participants in focus group 1 who were satisfied with the level of advertising on YouTube. G1P2 did not feel overwhelmed with advertising on YouTube. "Personally, the frequency of ads on YouTube when I am watching, I am OK with that." On the other hand, there were participants who appreciated **advertising that was entertaining** and were tolerant of longer advertisements if they were entertaining. G1P5 said, "The kind of content that marketers are creating nowadays is very eye catchy and we would not like to skip those video ads as well. I sometimes like those ads." G5P1 would tolerate even a one-minute advertisement: "If it is very catchy, I don't mind a 30 seconder or a 1-minute ad also."

DI2 F23 Delhi liked advertising when it contained a heart-warming message; this attitude was extended to the brands that she liked as well:

> I am willing to spend a couple of minutes on an ad if it is that good. So, at the top of my head, I can really think of the ads that have really struck me. So, there was this ad about a mom buying some swing for her kid. It was really made very nicely. There is an ad that talks about, I think it was a Nike ad which went really viral because of how it portrayed women. (DI2 F23)

Some participants liked video advertising on platforms like YouTube and were open to watching advertisements that had a connection with them. The consumers were willing to watch advertisements that were longer than 30 seconds if they had a relatable storyline. Some of the participants like DI 2 F23 were able to recall story arch from video advertisements that they liked. In focus group 5 there some participants understood that if they were not willing to pay subscription fees for the ad-free version of YouTube, advertising was unavoidable. G5P2 said, “On YouTube we do not have an option until you subscribe to the YouTube plus thing.” The ability of video advertising to emotionally connect with the target audience seems to cut through the clutter or resistance to forced advertising.

Einstein (2017) described the importance of context for the mass customization of advertising messages. Context along with postdemographic definitions of audiences helped advertising systems to predicting their personal situations. According to Einstein (2017), context provides information about the position of individual customers in their purchase journey and personalised advertising could then be targeted to cohorts of individuals who are at a certain stage of their purchase journey. Carah and Brodmerkel (2020) wrote about contextual advertising as an active ingredient in the user experience of a consumer driven by postdemographic audience definitions and rich contextual information. Participants had an affirmative approach towards the contextual information as they related it with personalised recommendations like related content, location-powered search results, and storage of their favourite webpages.

The ultimate objective of advertising technology is to change consumer behaviour by offering incentives for timely action or by building a long-term bond with the consumer. It is important to note that not all data collection is without consent. During the research, multiple examples came up where the participants were satisfied when the usage data that was collected by the digital platforms was used to create provide useful information like navigation and traffic information. The research also found cases of participatory surveillance; DI1 M25 Delhi enjoys the feature of automatic tracking and reporting in Google Maps – in other words, a “voluntary panopticon”.

In the previous section (User profiling and customised advertising) industry experts 1, 3 and 5 described the process of gathering digital usage signals of consumers and using them to segment and target the consumers for customised advertising. These processes are like processes described by Elmer (2003), Mayer and Cukier (2013), McStay (2018) and Szulc (2018) and to some extent resemble “everyware” as described by Kitchin and Dodge (2011). Szulc (2018) described how social networking sites developed detailed profiles of its users without their explicit knowledge. There was a general sense that the digital platforms like Google and Facebook knew a

lot about the participants and all the participants were able to list some form of data that is getting collected and analysed.

Foucault (2008) and his discussion of the mind experiment, “Bentham’s Panopticon” is widely used as the basis of examining digital surveillance. Mayer (2009) talks about digital data trails left by consumers with no control over how long these traces would exist. Gandy (1989) described the use of information technology and analysts deriving insights from multiple sources of data about individual behaviour as creating a digital panopticon, where the consumer is under continuous monitoring by nameless and faceless analysts. The data analysis creates a cohort of individuals having a similar disposition towards a particular product category and this classification makes them attractive for advertisers in that category. Campbell and Carlson (2002) make a very important distinction between a panopticon and advertising technology, where the sheer number of existing digital platforms prevents them from having a panopticon kind of control over individuals. Saulles and Horner (2011) coined the term “portable panopticon” for the mobile phone as it went everywhere with the user and was creating an observable data trail for digital platforms to collect and generate interest-based cohorts for selling to advertisers.

Bossewitch and Sinnreich (2013) bring in the concept of “sousveillance” where there the flow of information is bidirectional as opposed to unidirectional surveillance in a typical panopticon. The findings bear this out as participants seems to know a lot more about the digital platforms that they use and have the capability to surveil them through self-audits of privacy information and controls. They expressed that until such time they themselves experienced an issue, they were agreeable with the digital platforms collecting their data. Participants believed in customised advertising being important information as it resulted in getting benefits like better discounts, information about competing models of television or new phones being launched. Participants had a favourable opinion about big technology platforms like Facebook and believed that big tech has a vested interest in keeping their data safe and to not exploit them.

As an answer to the question that was posed in the literature review about how Indian consumers view digital platforms, participants have an affirmative approach towards them. Comparing the process depicted by the Industry Experts with the functioning of a panopticon, there is a degree of similarity in the possibility of continuous observation of an individual as they consume different digital platforms. The control exerted by an observer over the observed is not so explicit in the case of advertising technology as described by industry experts. There is emphasis on manipulating the behaviour of subjects in a panopticon by the mere act of observation. This has some similarity to the process of advertising where a target consumer is nudged towards a

particular communication objective; however, this action is very direct with an identified sender shown prominently in the advertisement.

As my research has shown, there is an elaborate system of collaboration between digital platforms, advertisers, data brokers and media agencies to create profiles of consumers for which data is shared voluntarily. Participants seem to enjoy when product features are created using this data like travel tracks, location-based “near me” search results, and most frequented places in ride sharing apps. The literature labels this as submitting to voluntary surveillance and calls this mechanism a technology-enabled panopticon.

These five themes emerge from the discussions above: 1) the likeability of advertising varies by platforms, 2) the ability of a platform to serve customised advertising impacted favourability, 3) advertising is seen as useful when it aids with tasks such as shopping, 4) advertising on digital platforms is more acceptable when the consumer has some control over the experience and 5) the quality of creativity played an important role in the likability of advertising. The quality of a creative aspect of video advertising was not limited to digital advertising and participants enjoyed such advertising on all media platforms.

There were some participants who did not agree with the five themes stated above and had their own mitigation strategies for digital advertising. There was a view of not clicking on any advertising. In focus group 2, G2P3 never clicked on any advertisement on any platform. “Talking about ads so I never click on an ad be it Instagram or Facebook or anywhere.”

Apart from not clicking on ads, there were some participants who did not like when advertising was customised. DI2 F23 was not comfortable to be placed in a cohort based on her demographics. DI2 F23 expressed that, “I like ads on TV because they are random. I don't like ads on my mobile phone or on my feed because most of the time they are customised.” This sentiment against customised advertising came up a few more times in other focus group discussions. The main opposition was against the very thought of an organization creating a profile based on past usage and/or age/gender classification.

A solution to advertising that could not be skipped is buying a premium ad-free service. G1P6 subscribed to the premium version of YouTube to avoid the hassle of skipping the ads every now and then:

> So recently they have launched a premium platform where you pay a certain amount, and all the ads are gone. So, I have opted for that because while you are listening to your favourite music and an ad pops and you cannot skip at times. It's a huge put off. (G1P6)

In this case, G1P6 was comfortable with paying a monetary price to switch off advertising and enjoying the video entertainment aspect of YouTube. Other participants in this focus group were not in favour of paying and one of the reasons not to pay was the usefulness of advertising to inform about various brand launches and choices available to them. **Advertising as useful information** came up across different contexts ranging from FOMO on deals to new launches of movies and tech products.

Advertising had a wide range of emotions associated with it and it is something that divided the participants across multiple dimensions. The responses ranged from total dislike for advertising to advertising as useful information. The first dimension was the context enveloping advertising. There were nuanced reactions to advertising, depending on the editorial context within which it was delivered. The most talked about contexts were advertising during information search, on social platforms and while accessing entertainment services. The second dimension was around the acceptance of customised advertising. The third dimension was compatibility of advertising with the native format of the digital platform. The fourth dimension was the ability of the consumer to control the advertising experience or the ease with which the advertising could be skipped. The theme of digital advertising as useful information came up spontaneously across different groups. The next section explores this theme in more detail.

6.3.1 Advertising as Information

As discussed in the previous section, digital advertising was seen as useful information when it was relevant to the topic being searched, researched, or watched. There were some examples where video advertising was liked if it provided information about the consumer's area of interest. One of the key categories that was mentioned was personal technology like mobile phones.

Some respondents in focus group 7 with male participants aged 25 to 34 had a favourable view of video advertising. Similar to other groups, there was an appreciation of advertising that was relevant to the topic being searched. Unlike other groups, they had a specific affinity towards video advertising. The key theme that came out was advertising as information. This is a slightly different variation of fear of missing out (FOMO). FOMO was expressed in both the female groups (Mumbai and Delhi Age 18 to 24) and was related to missing out on discussions on social platforms and information about shopping deals on e-commerce platforms. In focus group 7, FOMO was about information on new products and technologies. G7P6, talking about ads for mobiles phones, categorised it as necessary information: "They are entertaining and informative so there is a new mobile from Vivo. It seems attractive, eye catching. We like to know what are

the features in this mobile, we should have a look." Another participant in the same focus group, G7P5, said something similar about video advertising: "On TV the ads create awareness. Whatever new products are launched we get the information from the ads only. What is happening in the world of technology and where we stand." In addition to advertising on television as being information, G7P5 also shared sentiments like those expressed in other groups about liking relevant ads:

> If I am getting ads that are related to the topic, I do not mind ads like that. If I am searching for a topic on Google, say for example I am searching for education and if there are banners about education, I also get information. (G7P5)

Focus group 7 provided insight into another aspect of digital advertising, **consumer affinities**. The discussions in the group centred around interest in personal technology. These consumer affinities together with favourability about relevant advertising during purposive surfing, brought up **affinity advertising.** Participants talked about affinity advertising that is based on both the interest of consumers and their recent search activities. Based on the discussions, consumers were positively inclined to this kind of intelligence in their quest for information about their favourite topics.

Relevance was not restricted to advertising while engaging with a particular site; there were consumers who even appreciated when relevant advertising followed them across websites and platforms:

> When I search some products on Nykaa and when I open Facebook, I see the notifications for the same products about discounts. This is very good as they are providing me information as per my choices. They are aware about my choices and when I have searched about lipstick, they are not showing me ads for diapers. They are showing ads as per my preference. Once I have made a purchase the ads stop. (G3P1)

G3P1 brings relevance and affinity together in this statement as she understands that this information request is being serviced across the internet with some form of intelligence. This intelligence driving financial rewards, like discounts and offers, was an added bonus.

Beer (2009) postulated that the users might try and use the algorithms to their advance engaging into a duel as to who (algorithm or user) is able to manage the "right profile". Digital platforms and advertising technology gathers and processes data about transactions and online behaviour of users. The participants of this research cited numerous examples of how they benefitted by relevant advertising that helped them get great deals from across the internet once they have expressed that they were in the market for a particular brand or product. There were instances

when participants expected highly customised advertising and information in return for providing data access to digital platforms.

This is another example of customised advertising being seen as useful information. If it provides financial benefits like discount codes or offers, it is appreciated even more. Other members in the focus group 3 were not happy with the lack of control over retargeted advertising specially after she had completed the purchase:

> Once I buy a product, ads of related products keep showing up. I do not have any control over that. They do not continue for a long time; they will come for four five days and after that they will stop. Once I give ratings the ads stop. (G3P5)

G3P5 did not have any problem when ads followed her across the internet, but she felt a lack of control (TINC) to stop the same advertising once she has completed her purchase. In her case once feedback or a rating is given to the product that is purchased, the advertising delivery system receives the acknowledgement that the purchase has happened, and it should stop targeting the customer. As all the digital media channels cannot identify the completion of a purchase journey, G3P5 had found a solution to inform the advertising delivery system of her completed purchase by providing post-purchase feedback in the form of rating. This can be seen as consumers finding ways and means of interacting with an invisible advertising technology that surrounds them and achieving their objectives. G3P1 also narrated a similar experience where the ads stopped following her once she provided post-purchase feedback to the website.

There were other participants who did not like customised advertising as they believed that advertising customisation is based on creating a profile using their data. G7P5 had strong negative views about advertisements on social media sites like Facebook and does not want customised ads targeting her:

> They have not directly shared my data, but they have created my profile with my online activity and my friends. They have matched this profile; they may not have shared the data directly. Advertiser has given an age group to which they want their advertising to be delivered. I don't want Facebook to do even that. (G7P5)

G7P5 highlights a very important distinction between customised advertisements on social media sites and those on search or ecommerce sites. This distinction was made across groups where participants found advertising as information when they got relevant offers, new launches in their preferred product categories, and brands offering additional information to searched products while surfing online. Offline follow-ups were found to be irritating and frustrating as there is no

option to opt out of non-stop calls once a consumer has made an online inquiry and shared their telephone number.

Srinivasan et al (2016) discussed the linkage of online activities of users to their position in the product purchase journey and the cost effectiveness of online platforms to drive offline sales. The importance of online advertising during information searches for product purchases was described in detail and is explored in section 4.4 and 6.3. Participants were very receptive to advertising that was complementary to their information search and across all aspects of their purchase journey. As Srinivasan et al (2016) found in their research, consumers were equally interested in information about FMCG brands and expensive items like phones or television sets. Some participants of my research were able to make the connection that their purchases of advertised brands drove revenues for digital media platforms, in other words they validated Smythe's (1981, p. 233), 'audience power'.

Carmi (2020) outlines the dystopian possibilities of advertising-funded digital platforms, whereas some participants in my research felt empowered by the level of recommendations and personalization, especially when they were on a path to purchase products. Some participants felt that there should be no personalization and preferred generic ads.

6.3.2 Source of Revenue for Digital Platforms

As discussed in chapter 1, advertising is the main source of revenue for popular digital platforms such as Google, Instagram, and Facebook. Section 4.4 describes the exchange of user data for subscription free access to services from AFDP. Advertising as the source of revenue was not understood by all the participants of my research. Marketing practices described in the previous section (6.3) exhibit how marketers can reach out to users at different stages of their product purchase journey. While the underlying digital advertising technology was not understood by some of the participants, they could recount several examples where they felt the connections between their actions on one platform and advertising on another. G4P5 recounted, "when we search for something on Google, advertisements for the same show up on Facebook and Instagram. Whatever we search on Google, its advertisement increases automatically". Another participant G4P6 explained, "for example, if you have searched for a product on Google or on YouTube, things related to this search, like offers or discounts will automatically show".

Participants imagined that there was interconnectedness not only between platforms owned by a single corporation but also across competing platforms. There was a sense that platforms across the internet were linked through an invisible layer that could be perceived but was difficult to understand. Some Everyday Users were able to explain how personal identifiable information (PII)

was used as a key to identify them across their online engagements. G4P6 identified telephone numbers as one such link: “Our registered phone number where the OTP comes, is with the service provider and this is the key to everything”. G4P1 described the IMEI number as another connector: “The device that you are using you can be tracked via your IMEI number”. These two participants identified their phone numbers and their mobile device IDs as two identifiers that track and unify their online behaviour. Some of the participants were unable to comprehend how their activities on one platform were triggering advertising on a different platform. Their surprise was vocalised by a similar expression: “How is that even possible?”.

There is abundant literature describing the workings of ad tech and the role played by tracking cookies (Elmer 2003). Wall (2006) outlined how cookies are used by DoubleClick to link user activities across the internet and create a detailed consumption profile. McStay (2017) produced a detailed list of identifiers used to track and target individuals across the internet. However, as seen above ordinary users could not grasp the entire gamut of identifiers that were used to link their online activities.

Focus groups 1 and 5 consisted of participants who had some experience of working in the marketing and communications industry. These participants had domain knowledge about media economics and advertising technology. Participants in these focus groups were able to identify the two sets of customers that digital platform served and how that manifested into two distinct choices for users of digital platforms. The users of digital platforms could either pay a subscription and access the advertising-free versions of some of the most popular entertainment apps like YouTube or be served advertising from the other customer of the digital platforms: the advertisers. The economic dependency of digital ecosystem on advertisers was understood and accepted in these two focus groups:

> In India everyone likes everything free so irrespective we get ads on our favourite application that we use. Out of 100, 99 people do not buy it the premium. So that is one of the reasons why our data is being shared with advertisers and we are getting ads. if we buy those applications that we are using and if we delete those applications that are of no use to us I think that will help us to protect the data which we do not want to go into the wrong hands. (G1P6)

There is an understanding that their usage data will be shared with advertisers, and they even expected responsive and targeting advertising that is useful information in their purchase journey:

> There is a thing on digital that is called programmatic that means that they trace the behaviour of that person, so wherever I am going they are following. If she has done any search about Maldives, the next things she will get is an ad about Maldives. (G5P5)

G5P2 described programmatic advertising: “It is like a subconscious memory that even internet has, it is tracing your search pattern or your behavioural pattern. Even if your location is on its tracking you”. Some of the participants in group 5, who were working in the media planning function of a Mumbai-based media agency, were able to illustrate how programmatic advertising can straddle the entire spectrum of online experiences and create behavioural affinities. These affinities are used to create the probability of a person being in the market for products and services to engage them with customised advertising. DI5 F32 expressed her POV on customisation: “Overall, there are pros and cons to everything but it's useful no I don't feel exploited. It's more of convenience and knowledge sharing and ease of use vs confidentiality of data or getting exploited or getting addicted”. As the data for programmatic advertising is collected across online engagements, it is largely invisible to the target audiences of the campaign.

DI5 F32 described multiple advertising options available with digital media platforms and the primacy of advertising as the main contributor to the income of platforms such as Google, Facebook, and LinkedIn. She worked in the marketing industry and was able to articulate the relationship between advertisers and digital platforms and the level of data sharing that enabled customised advertising.

> Facebook paid ads are there they have so many groups where people post ads, or there are B2C companies that do a lot of marketing also. Google PPC ads are there, paid ads are there, when you search for anything the search engine optimisation will give you organic results but there is a box in which the first search will appear and that is paid. Google PPC ads are costlier Facebook PPC ads are also there and LinkedIn PPC ads are also there. PPC is pay per click. (DI5 F32)

Being a practitioner of the marketing industry, DI5 F32 expressed familiarity and comfort with the data collection practices of digital platforms. She used the internet to her advantage and was confident of her online practices to minimize any adverse impact.

However, in other focus groups, participants were unable to comprehend how their activities on one platform were triggering advertising on a different platform. Their surprise was vocalised by a similar expression: “How is that even possible?” In these groups there was a lack of shared understanding of revenue model of digital platforms. Some of the main business models that

were discussed were 1) advertising, 2) selling user data, 3) commission on sales from brands, 4) subscriptions and 5) revenue from internet service providers. Topics 1 and 2 were referenced most often, although the engagement on the topics amongst participants was very low as compared to the other themes such as privacy concerns, data awareness and data control.

Some participants in these focus groups were able to outline how advertising was driving the revenues for digital platforms:

> This is confirmed that they make money through advertising. When we watch an ad we get attracted to the product. Say for example when we see an ad for beauty products, the first thing that we do is search for it. Then we order the product, so the money is coming from ads only. (G3P4)

G3P4 was able to describe the cycle where purchases made after watching an ad is what drives the digital media ecosystem.

There were other participants who imagined that advertising plus the sale of their data to advertisers was the main revenue source for digital platforms. DI2 F23 said “I think the revenue generation model comes from advertisement probably selling the data to the companies without consumer's consent”. G4P3 commented about Cambridge Analytica and Facebook: “Openly no one is selling our data, but they might be doing behind the back. Facebook has got stuck badly in this”.

DI2 F23 articulated that advertising and sale of consumer data were two main revenue streams. G4P3 and G4P5 expressed a similar understanding that digital platforms were selling their data and made a reference to Cambridge Analytica to prove his point. Source of revenue for broadcast media was also not very clear to some of the participants:

> In the TV world they say that when the programs get more ratings (TRPs) the channels get more money. I do not understand when we watch TV how do the TV channels make money? I think when we do data recharge, they get money at that time or maybe advertising also. (G2P1)

G2P1 is describing the understanding that some of the participants had about the business model of television as a medium and the relationship that TRPs have with advertising rates. Some other participants found it difficult to show a similar understanding for digital platforms.

Some participants felt that the digital platforms generated revenues when the users paid for their internet connections. G3P2 could not understand how the money reached the digital platforms: “We do not know how the money goes to the platforms?”

G2P3 shared her opinion that Facebook is in the business of selling her personal details like phone numbers:

> Sometime back there was a leak in Facebook where they shared all the private information people had, mobile numbers and everything so there was leak from their end. We do not know if there was a leak or actually, they wanted to sell the information. It's bad obviously. (G2P3)

A fear of data leaks and the Cambridge Analytica scandal convinced her that Facebook was in the business of personal data like mobile numbers to make money. Other participants outlined factors like Cambridge Analytica and receiving messages and calls from random brands as the main reasons to believe that their personal information was being sold. A lack of detailed discussions in the focus groups about the business model of digital platforms, the opaque nature of digital advertising technology and scandals like Cambridge Analytica, contributed to the belief that selling data was a source of revenue for digital platforms. To emphasise the point that these participants did not mean that Facebook was selling their app usage data to advertising, they felt that their personal details like phone numbers were being sold to advertisers. This resulted in them receiving so many random calls from companies that they had never interacted with.

Some participants in these two groups could describe the business model where advertisers subsidize the users of certain digital platforms. Some of the applications had premium ad-free versions and examples of YouTube Premium, Hotstar and Saavan were mentioned throughout the research. Participants appreciated customised advertising and recounted examples of personalised topical advertising:

> I went to Bali recently. I was doing a couple of searches on what are the places to go what are the things to do. Once I do a search, tomorrow again I do a search, automatically I start getting information which I wanted at one point of time, and I still require. So consciously subconsciously I actually might not need it, but I still have a want for information or brands that are being pushed out to me. (G1P6)

G1P6 favoured a very high level of customization and plots a clear path from giving signals to advertising engines of search apps and contextual advertising served proactively:

> I just want to add one thing to it it's that in the current scenario you are searching for things, and somebody is getting that data where you are actively interacting with the portal, a platform whatever and the data is being taken up by that platform and given to somebody else he's targeting that's fine. (G1P2)

Shoenberger (2017) highlighted the issue of user control over the type of data that is used in generating highly personalised advertising and participants felt a similar lack of control over the customization experience. Some of the participants were thrilled after the Facebook ad preference audit as they could switch off the customization but were a bit disappointed that they could not find a similar solution for other applications.

Some participants spontaneously described experiences with online advertising that made them extremely concerned about being surveilled. There were instances where participants attributed the accuracy of advertising personalisation to the ability of the application to listen to their conversations. This practice of advertising customization being driven by app like Facebook listening to their conversation was described as audio snooping. Audio snooping was spontaneously recounted by at least one participant across multiple focus groups. The next section delves into the attributes of audio, video and spatial surveillance that emerged from experiences and trepidation of the participants.

6.4 Audio, Video and Locational Surveillance

Whitaker (2000) outlined the importance of information profiles to productise the audience. The process of datafication by advertising technology using cookies has been described in detail by scholars such as Elmer (2003), Wall (2006), Van Dijck (2013). Glass and Callahan (2014) described this as data muscle, Lyon (2014) stressed the importance of metadata in creating data doubles (McStay 2017), whereas Mayernik and Acker (2018) highlighted the opacity of metadata. Carmi (2020) described the invisibility of advertising tech to ordinary users of AFDP. This invisibility drives participants perception of ad tech surveiling their personal situation that can be felt but never fully understood. This is the most intriguing and more intrusive side of advertising technology:

> The problem when this data accumulation gets into a privacy mode and becomes snooping. Audio snooping today has become a thing for example if I have my phone on and I'm talking about something and I've personally experienced it, that two days later I get served an ad on that. I have not even searched it, but we have had discussion around something. That is scary. (G1P2)

My research participants described audio snooping as apps such as Facebook and Instagram, listening to conversation in the vicinity of a mobile phone and serving advertisements related to the topic of conversation. This issue came up across multiple groups in both Mumbai and Delhi. G1P2 referred to this highly intrusive example of customised advertising and as a response two other participants in the focus group recalled that they had heard about something similar.

Though G1P2 was the only one in the group who had encountered audio snooping first hand, some other participants were in strong agreement that it was happening.

Participants in other interviews shared similar examples of receiving advertising triggered by conversations in the vicinity of their mobile phone. The use of audio snooping for personalised advertising was not mentioned by any of the Industry Experts interviewed for this study. In focus group 5 in Mumbai, G5P1 gave an example of personalised advertising she could receive based on the topics that were discussed in the proximity of her phone. “I will not be surprised now if I get an ad pop up saying that are you tired of seeing advertisements. Please download this app and it will block all that”.

Multiple participants in focus group 5 agreed with this theory. Some participants were unable to fathom how customised advertising could address their immediate personal situation so accurately. They attributed the intimacy of advertising to audio snooping. In focus group 5 there was a chorus of approval and examples of ads that were served after discussing vacation spots. There was a strong feeling that their phone was listening to everything, and the words were being used to customise advertising. Generally, participants recounted that they encountered such advertising on Gmail, Facebook, and Instagram.

> What bothers me a lot is if I am thinking about something or talking about something, and the same ad pops up. Every time I am on Facebook, and we are speaking about some place where we want to go, and we literally see an ad. (G5P4)

G5P4 shared an example about discussing a medical condition and receiving ads about it. “That day I was talking with someone about PCOD, and I got ads about apps that give information about PCOD I also got some ads information about it. You cannot even talk in privacy.” These extracts represent multiple instances of participants discussing a topic in the proximity of their mobile phone and receiving advertisements about the very same topic within minutes to a few days of the conversation. The chronology of events made them question everything they know about advertising technology and build their convictions about conversations triggering ads on Facebook.

G5P2 shared an email that she received while the focus group participants were discussing travel related ads: “While we were talking here, I got an offer from Makemytrip for three nights four day.” G5P4 was expecting advertising related to the topic of discussion, though she did not encounter an advertisement while the discussion was ongoing. “We will all get messages, and I will get about somewhere and she will get about Vietnam.”

Most of the participants in focus group 5 expressed a strong conviction about audio snooping for personalised digital advertising; however, the reaction was that of amusement rather than fear or stepping back from engaging with digital platforms.

A participant in focus group 5 tried to explain the level of data intimacy that was possible using programmatic advertising (see programmatic), but others were convinced about audio snooping. It seemed participants adopted a folk theory to make sense of a complex ecosystem.

The folk theory that mobile phones were listening to conversation also came up in in-depth interview 2 in Delhi.

> The next time when I'm having a candid discussion and my phone is in my pocket, I'm talking about something that might be very personal to me and if I see something of that sort, now imagine me discussing pregnancy of a friend and my mobile phone, I've taken a screenshot also, said that for pregnancy in Gurgaon you need to visit these these hospitals. We were discussing the pregnancy of one of our friends that she is carrying a baby and I felt offended and felt violated. (DI2 F23)

DI2 F23 shared an example where she was offended by the fact that her phone was listening to her conversations to serve advertising. As the information about hospitals came up after her discussion with her friend, it was strong empirical evidence of causality. Apart from specific examples as described above, there was a general sense that apps were listening to the conversations:

> I feel like when I am talking about something an ad can come from Facebook and Google about that very same thing. I don't know how. At times I feel that I get advertising of brands or products about which we are only talking. It feels like they are listening. I am not sure; I believe that is not the case. At times it feels like they are listening. (DI1 M25)

This spontaneous reference to the folk theory was made while the participant was discussing the business models of free applications like Facebook, Instagram, and Google.

Skeggs and Yuill (2016 a, p. 391) “found that the density of network influenced how robust Facebook algorithms were”. As the participants of my research experienced the intimate accuracy of the Facebook ad delivery algorithm, they might have a dense and active network. However, this assumption was not validated by an audit during the focus group discussion. The density of Facebook network for my participants was not a recruitment criterion and is assumed based on the level of discussion about Facebook usage and the spread of their connections. The invisibility of advertising technology was uncovered by Skeggs and Yuill (2016b), when they discovered that

Facebook ad targeting was enriched by off platform user data. Without a toolkit to guide my participants about the source of accuracy of advertising messages, they believed that the **ads were listening to them**.

Apart from audio snooping, some participants expressed their fear about the use of the mobile or laptop camera for video surveillance. None of the participants had experienced it, but they believed that there was a strong possibility that it was feasible. The participants in group 2 also felt that hackers could be listening to their conversations using microphones. A few of the participants were worried that hackers could also watch them using mobile or desktop cameras. They were not too worried about surveillance for customised advertising, but more worried about various apps asking permissions to access the camera roll, microphone, and camera:

> I don't know how true it is but then any app developer or a software engineer or somebody they can hack into it and if they want, they could listen to this very conversation just through the media and mobile that we have right now. They can watch us through the cameras, they can listen to us, that's bad. (G2P3)

G2P3 imagined that it is possible for hackers to listen and watch them using camera and microphone.

G2P3 was also worried about the possibility that hackers could use the permissions granted to various apps to listen to their conversations and watch them. She had not experienced such a hack on her devices but was not sure if she will even find out that she has been hacked in this manner.

> They take permissions for microphones on our device what are they accessing through it? Asking for camera privileges also its not photo roll but camera privileges what do they want this for. I don't know how true it is but then any app developer or a software engineer or somebody they can hack into it and if they want, they could listen to this very conversation just through the media and mobile that we have right now. They can watch us through the cameras, they can listen to us, that's bad. Like why? (G2P3)

G2P1 was convinced that this type of hack was happening a lot, though both did not have any examples of their phones being hacked. "I second with her, that's true it happens a lot." The folk theory that there are a lot of hackers out there who will steal data also came up in focus group 3 in Delhi. G3P2 described helping her mother-in-law when she got hacked: "There are a lot of hackers out there. It happened to my mother-in-law. My mother-in-law had gone to Australia and her Facebook got hacked".

Similar examples were recounted by G4P4 who provided details about being hacked and his credit card data being misused. He also stressed that platforms like Facebook were not safe and G4P1 added his understanding that other platforms were getting hacked regularly. G4P4 said, “You know Facebook get hacks five times a day and no one knows. There are bounty hackers who are deliberately hacking Facebook and getting paid for it. There are those vulnerabilities that they want to plug”. G4P1 contributed to this discussion with his point of view that, “Not only Facebook, but WhatsApp and Twitter is also getting hacked. We are seeing movies where there are people who do certain work and are caught. Data stealing is happening”. Movies and television dramas also play an important part in how technology is perceived and contribute to the folk theory. G4P1 expressed that the main source of his understanding of hacking was from the movies.

The topic of excess monitoring had references to Cambridge Analytica, as well. A few participants had a very evolved understanding of the importance of their data and how it is being used for things more important than advertising. G1P2 talked about the use of consumer data to influence elections and democracy. The participant felt that this could be the reason why digital platforms were in the crosshair of regulators: “Things like Cambridge Analytica etc. these are being used to build government. Which have a far bigger impact on people's lives than advertising.” Customised advertising was not considered a troubling example of surveillance, but the participants were worried about the use of advertising infrastructure to deliver impact on political choices.

locational surveillance emerged as another cause of concern for some of the participants. Though the instances of this popping up in discussions were fewer as compared to the topic of audio snooping:

> I have noticed in Uber that where you go the most, yes, it is easy that they are saving your location, but it also means that it is recording that this girl goes to this place lots of times. So, its saving all our history. (G2P5)

Another participant in focus group 2 had a toolkit (see section 5.6 privacy red lines) to manage her privacy settings and she disabled geotagging on her photographs to avoid people surveilling her location:

> Uber and Ola track when you're going from point a to point b basically those rights are there are you going from one location to another location every day. people will know this is your house this is your office people would know this is a residential area. (DI 5 F32)

DI5 F32 was aware that her regular use of ride sharing apps can easily reveal her home and office locations; still, she preferred digitally mediated ride hailing service to the physical hailing of a cab.

In this section participants explained the surveillance experiences encountered by them in their daily lives. Audio snooping for serving personalised advertising by AFDP such as Gmail, Instagram and Facebook was described by multiple participants. Some participants described their anxieties of hackers taking control of their video cameras or listen to their conversations. Video surveillance was mentioned as a shared anxiety, though, none of the participants had any first-hand experience of a phone being hacked for this purpose. Locational surveillance was mentioned as a possibility as ride sharing companies had records of their movements but none of the participants had a corrosive experience with location tracking. Next section builds on these data experiences, collate them into folk theories and compare them with existing scholarship about this topic.

6.5 Folk Theories of Data Gathering

In the previous sections, Industry Experts were able to outline the process of collecting detailed behavioural and transactional data about consumers at different stages of their purchase journey. Some participants were able to describe the revenue model of digital platforms, while others resorted to a common-sense understanding of how these platforms generated revenue. As documented in previous sections, concerns about audio snooping came up across multiple groups and participants imagined that their mobile phones were listening to their conversation to gather topics to serve personalised advertising. In this section the genesis and structure of folk theories mentioned in the abovementioned sections research will be discussed and compared to scholarship in this field. At the very heart of data gathering is a complex information technology ecosystem powered by a plethora of information in the digital exhaust of users aptly named “Infoglut” by Andrejevic (2013)

Eslami et al (2016) used folk theory to explore user perception of the process of curation of Facebook feed. Toff and Nielsen (2018) used folk theories to understand the access to news from non-consumers of traditional news sources. Mayer-Schönberger and Cukier (2013) described the predictive power of algorithms and its accuracy in providing answers even before a question is sought from a technology platform. DeVito et al (2018) described the role of “endogenous information” and information from “exogenous sources” that combine to help people form folk theories.

India has been a mobile first country where all the information is available via a handheld device, users try and find simplistic answers to the complex digital environments that they encounter. Some common-sense theories about data gathering practices of digital platforms emerged from

various discussions across gender, age groups and cities. Some of these theories were about a specific app and others were about personalised advertising. The most commonplace theory was the folk theory of customised advertising.

6.5.1 Folk Theory of Customised Advertising

There were spontaneous references to the topic of receiving advertising related to candid conversations in the vicinity of their mobile phone. Participants recounted numerous examples of receiving advertising that was intimate to their current circumstances. These circumstances were related to their medical condition, travel plans and shopping preferences among others. Participants imagined causality in a chronological occurrence of events. This theme of audio snooping or audio grab came up across multiple groups with participants citing personal experiences and word of mouth examples.

In focus group number 5 multiple participants agreed to this theory. Despite working in digital media planning, they were unable to fathom how their personal thoughts or discussions could trigger personalised advertising, with audio snooping being the only explanation. Throughout this research this theme of “ads are listening to me” was substantiated with examples about medical conditions, holidays and at times a general feeling that was not backed by empirical evidence. In this focus group a participant explained how programmatic advertising predicts user preferences. Industry Experts were able to provide a detailed set of practices that is used to create highly customised audiences in a manner described by Mayer-Schönberger and Cukier (2013), but the participants from both the sets described above understood chronology as causality. The resultant attitude of participants who espoused the folk theory of customised advertising was quite similar to the one reported by Ytre-Arne and Moe (2020). Participants in this study were obliquely aware of the existence of algorithms and a heightened sense of resignation about the customised advertising as something they have to accept as part of accessing software enabled services.

Widespread adoption of this theory as validated by its occurrence in multiple interviews across cities exposes a limited understanding of advertising technology and the predictive abilities of algorithms. As witnessed in focus group 5, it was impossible for the participants to fathom the intimacy of advertising on platforms like Facebook except for the theory they developed. The level of discussion and agreement amongst most of the participants was evidence that the absence of “endogenous information” of the Facebook platform, they had relied on “exogenous sources” and their heuristics to develop this theory.

6.5.2 Hackers Can Easily Access My Data

There was a sense amongst the participants that they did not fully understand the mechanism of how computers or phones are hacked. There was a conviction that it happens a lot and this had something to do with the permissions that various applications seek when they are installed. The hypothesis that there are hackers out there was discussed in focus group 2, 3 and 4 and in all instances, there was general support from other participants. A growing push about transparency on which app is using which feature on the smartphone gives the users a glimpse into what all is possible. This newfound knowledge is used in combination with news information about data leaks to build folk theories that there are “hackers everywhere” and they can somehow gain access to their cameras and microphones and “listen and watch us”. Rip (2006) wrote about folk theories that try to pass as empirical evidence and in this case a lot of participants referenced hackers as seen in movies to justify the existence of hackers being everywhere with the ability to switch on their cameras and microphones. Eslami et al (2016) conducted laboratory experiments with multiple participants to understand and reported that different participants arrived at similar folk theories about the curation of social media feeds. The case of hackers was similar with three different focus groups which discussed the prevalence of hackers around them.

6.5.3 Data usage by TrueCaller

The third folk theory was about TrueCaller's ability to identify all incoming callers. As discussed in section 5.3and 5.5, the foundation for this theory was when TrueCaller is installed it asks for permission to access the contact list of the user. Participants believed that this contact list is compared with contact lists of other users to build a global database of callers. This led to anxiety amongst participants that their data was indirectly being put in the public domain and also triggered another assertion that TrueCaller was selling their phone numbers to other businesses. This is a simplistic understanding developed using heuristics and was repeated across multiple groups.

6.6 Conclusion.

Marketing practices for creating customised advertising can be described as the process of converting business questions into marketing objectives to curate a target audience from the universe of users. Data integration helped in identifying the right set of digital platforms that could deliver the target audience. The end goal of an advertising technology system is to place the right message in front of the right audience for the duration of the campaign. As discussed above, most of the mobile phones in India work on Android operating system that allows the use of third-

party cookies, enabling the advertising technology to track users across the internet. This interconnected layer of technology is invisible to everyday user. Advertising surveillance has reached the level of sophistication required to profile users into interest and affinity groups. Consent for the use of SDKs owned by data brokers along with third party cookies is buried in the lengthy EULA that some participants in this study had never read. This lack of explicit informed consent creates an aura of intrigue around customised advertising.

Industry Experts were able to describe a complex array of legal practices that aggregate identities across a large swath of internet activities to create qualified audiences for branded communication. Some Everyday Users were not privy to this intelligence and for them highly accurate personalised advertising delivered by apps like Facebook, Instagram, woven together with digital platforms asking for permissions to use the microphones, pointed towards audio snooping. My research explores the concept of "algorithm awareness" (Bucher 2017) and makes a unique contribution in understanding how Everyday Users experience algorithms through personalised advertising on AFDP.

Willing participation in providing extensive usage data to a digital platform as a pre-condition of access is one form of participatory surveillance. Participatory surveillance is also used to describe the use of technological platforms to surveil health, exercise routines, or food habits. Bossewitch and Sinnreich (2013) called this practice self-surveillance and presented an affirmative use of digital technology. It is also called "Quantified Self" where people use a range of technology devices to create a quantified image of their existence over which they have complete control. Ganascia (2010) introduced the concept of "sousveillance" where the many could watch a few, especially people in power and public service. The participants felt that the internet had a "subconscious memory", and a lot of locational features were helping them stay safe while they navigate Indian metros and provide them a sense of familiarity in unknown places. These elements of advertising technology that are used to provide customization like location and history of online behavioural data, are seen as enabling by the participants.

Participants were positively inclined towards voluntary surveillance, especially by Google Maps. In this instance, relevance made a big difference as to how the use of location is viewed. Use of location for directions and ride sharing is viewed favourably, whereas the use of location by Facebook to broadcast their position to friends is deemed unnecessary. The fear of being left out as described by Whitaker (2000) is a powerful motivator for people to accept all the terms and conditions of mobile apps as captured by the term FOMO expressed in multiple groups.

Data usage practices in India mirrored the ones defined by scholars (Kitchin and Dodge 2011, McStay 2017); however, some participants of this research exhibited a pragmatic approach to

personalised recommendations. They had an affirmative view towards voluntary participation in the exchange of convenience for privacy if the services were free and unique. This view did not inhibit the participants to try and understand the invisible layer of technology. They theorised about the different possibilities about how their preferences and intent relate to their technology-mediated experiences.

Lack of awareness about business models, increased visibility over hardware access by digital platforms, and enhanced accuracy in predicting personal situation of participants led them to postulate folk theories of data gathering. Three folk theories that were mentioned across multiple interviews were: 1) Folk theory of customised advertising where participants imagined that their phone was listening to their conversations and believed chronological occurrence of events as evidence of causality, 2) hackers can access their data easily and 3) data practices of TrueCaller. Scholars have outlined folk theories about information technology (Rip 2006), distributed discovery (Toff and Nielsen 2018), journalism (Palmer 2019), curation of social media feeds (Eslami et al. 2016), local newspapers (Nielsen 2016) and algorithms (Ytre-Arne and Moe 2020).

Chapter 7 Conclusions

7.1 Introduction

This chapter returns to the Research Questions.

- RQ1: How do young Indian smartphone users experience personal data and its reciprocal use by digital platforms to provide access to software services? And
- RQ2: How are audiences for customised advertising created from digital data trails of users?

India has witnessed an unprecedented adoption of digital platforms and a lack of regulation has created a pool of internet users that have become smartphone literate through self-learning. According to my research, this literacy was limited to the functioning of the device. This research aims to understand the process of generating targeting information from data traces of users and their data experiences while using software-based services via their mobile devices. Semi-structured qualitative interviews were used to understand the processes involved in structuring advertising surveillance and the experiences of participants when encountering personalised advertising.

As these services were considered “free” there was a sense of obligation to grant access to personal data and digital platforms which seems to be following the “benefactor before beggar” strategy as described by Cialdini (2007, p. 24). Some participants felt a sense of resignation that they were insignificant as compared to big technology firms. Size, a factor that made big technology companies trustworthy, worked against them when it came to privacy controls. Some participants expressed awareness that the answers about the use of their data for customization and privacy policies were covered in EULA. Safety and the large number of users around the participant was linked to FOMO. If most of the people around them were using an app, it gave the participants a sense of safety, but it also inculcated a sense of missing out if they did not follow what everyone around them was doing.

Another reason for participant apathy was the trust built over years of usage and a sense of safety in numbers. The trust threshold was an individual construct, but there were some common heuristics that emerged across interviews. Apps were installed after a sense check that combined a check on the app store, level of permissions, track record, social acceptance, and importance of the problem that app is solving. Once any app has passed the trust threshold and has been installed, it was considered safe.

In the interviews, the discussions about privacy were subdued indicating a low level of historical interest in the topic, but the resultant explorations brought out the topics that the participants considered sensitive. One of the most sensitive topics was access to cameras and microphones. These privacy red lines were strongly expressed by female participants both in Delhi and Mumbai along with access and visibility of their photos. An equally sensitive data type was access to financial information and granting access to bank or credit card details which was a major threshold to cross.

Participants happily recounted their use of advertising technology to help them accumulate useful information about discount offers from e-commerce sites and new product launches from their set of advertisers. This activity was undertaken without the explicit knowledge of ad tech, but through interacting with the visible part of the ad serving ecosystem, customised advertising. Some participants expected personalised recommendations and advertising as their right after parting with copious amount of data. The participants understood that “An Advertising Panopticon Rewards Participation”**.**

The next three sections provide a summary of this thesis along the three themes: Data, Privacy and Surveillance. These sections are followed by key research contributions, implications, limitations, and future research areas.

7.2 Data

Data experiences of participants were driven by the trusted status of digital platforms. They believed that universal access to mobile internet has empowered them through social connections, a continuous stream of news and information, enhanced mobility options, improved control over their data choices and surveillance of their social and economic environment. The affirmative view of digital technology was built on trust. A bond of trust had formed between digital platforms and the participants, cemented by a long trouble-free association. Trust thresholds were built on complex interactions between users and digital platforms over a period. Clearing the threshold resulted in rewards for platforms like increased sharing of sensitive data and the ability to cross privacy red lines. A trusted application could collect data and deliver personalised services until such time a trust deficit emerges.

Apart from tenure, trust also emanated from the corporate ownership of the app, widespread usage in society or close social circles, a unique service offering, fairness of data permissions, reviews, consumer ranking and convenience provided by the digital platform. Trust also enables the exchange between usage data and access to software-based services offered by digital platforms. The discussion around trust can be summed up as: **Trust Breeds Compliance**.

Participants understood that they were exchanging their usage data for the convenience of accessing the services of digital platforms. Participants imagined that they had a one-to-one relationship with digital platforms and could not fathom the connections enabled by advertising technology. This interconnectedness enabled platforms owned by competing organizations to share the same spine of identifiers to track users across the internet. As the research method included interviews with Industry Experts, it benefitted from a preview of contemporary industry practices in this area.

Data exchange that fuels the economics of advertising-funded digital platforms was viewed as fair, as the participants valued free access to world class technology and data-enabled product features. The ability to interact with cutting edge technology thrilled some of the participants. Content recommendation algorithms were being used to stay informed about the latest developments in the field of personal technology, location services were used to search “nearby” and e-commerce recommendations and cross platform tracking was repurposed to get the best discount offers for favoured brands. Users in India have generally embraced the role of mobile internet in their daily lives and in some instances extracted rewards from the exchange by using advertising technology to do their bidding.

These hacks helped some participants to observe their favourite brands through the use of ad tech. These hacks did not empower the users to understand the synthesis of advertising technology, but only helped them to utilize the visible elements for desired results. Participants had a sense that by sharing their data they were entitled to reciprocation from the platforms. The reciprocal relationship can be termed as participatory panoptic sort, where users agree to be observed and in return use the digital platforms to sort the complex information environment. This reversal of the panopticon from a traditional one-to-many observation, to an inverted panopticon of many observers-to-one (the advertiser) was evident throughout the research.

This simplistic view of a complex digital ecosystem made it difficult for Everyday Users to grasp the challenges related to data usage and privacy. Table 2 list the tactics deployed by the participants to mitigate their disquiet about lack of control over their data experiences. Knowledge acquired from news and other exogeneous sources had made them aware of the risks of sharing their financial information online, but the gaps in literacy about data usage and data privacy resulted in a privacy blind spot. In the research, female participants were specifically careful about the privacy of their personal information like their home address and photos. At the same time, there was a pragmatic approach with regards to sharing personal data when it was required for accessing location-based services and e-commerce deliveries amongst others.

Through this analysis, six main reasons emerged for not investing time in understanding EULA: 1) the safety of data with large corporations, 2) the safety in large number of users, 3) a fear of missing out (FOMO) on popular apps, 4) there is no time (TINT) to read the lengthy terms and conditions and 5) there is no option (TINO) as there was no alternative to the apps, 6) there is no control (TINC). FOMO was the most powerful reason for the continuous patronage of digital platforms.

Some participants expressed frustration with the lack of regulations about the safety of their online data and demanded laws to safeguard their right in cyberspace (see regulator). The Personal Data Protection Bill (Indian Parliament, 2019) had been pending before the Indian parliament for a few years before it was withdrawn in 2022.

7.3 Privacy

Participants highlighted the benefits of exchanging privacy for convenience. They were able to express their requirements of data privacy but also expressed their inability to control this aspect of their online experiences. Participants listed their sensitive information that they wanted to protect, and financial information came up as the most sensitive piece of data. Some participants were worried about their browsing histories being used for better customisation as this could result in inadvertently revealing their private information to family members as mobile phones are shared with family members in India.

Overall, privacy measures were relegated to the management of the visibility of social media posts, especially photos, and watch, search, and location history. Lack of context or a prescribed framework to understand the implications of privacy controls led the participants to restore to a common-sense understanding in interfacing with highly predictive advertising technology and advanced control dashboards. As described in the introduction, unprecedented growth of mobile internet with little regulatory oversight has left the users in India without much historical understanding related to personal data privacy. This combined with the fact that in India mobile devices can be accessed by other members of the family creates privacy challenges that are unique to India. A mobile device that is social and not strictly personal, can reveal a lot about user's interests and preferences through algorithmic recommendations.

Scholars (Gandy 1993, Davies 1997, Turow et al 2015) wrote about the computerization of databases and enhanced speed of computing and expressed their disquiet about how it enabled various sections of the economy like finance and retail to generate detailed consumer profiles and accelerated datafication of the population. Historically, India has been a cash economy and users experience with big data providers such as credit rating agencies or other credit systems was very

limited. The sense of being surveilled came up due to the feeling that ads were following them across the internet. The Cambridge Analytica scandal, increasingly visible permissions from apps and a lack of understanding of business models made the participants uncomfortable about the source of accuracy of personalised recommendation.

The act of providing consent brings the issues of data privacy into the everyday consciousness of the audiences and can help them develop countering mechanisms. These mechanisms have the ability of being developed collectively and deployed individually to have an exponential impact. Two key deterrents were identified in scholarship (Gurses and van Hoboken 2017, Sadowski 2020) that prevented the participation of individuals in the meaningful management of privacy controls: obfuscating EULA and the rapid pace of service enhancements. However, the rapid pace of service enhancements that are seen as a deterrent to privacy management in scholarship were viewed as great features by participants of my research.

Privacy campaigners and regulatory pressures in the Western world have forced digital platforms to provide more visibility to the users about profiling and advertiser interests. Digital platforms are providing increasing visibility to privacy policies and control over data experiences to their users. However, there exists a lacuna in the communications surrounding this enhanced control and privacy empowerment. In the absence of motivation and lack of communication from digital platforms, end users are left to their vernacular understanding of personalised advertising and other data experiences. The self-audit of ad preferences on the Facebook mobile app revealed enhanced personalization controls that were available to the participants. They were also able to examine their algorithmic profile and were able to experience the output of algorithmic sorting. This opened the possibility of interested audience members to garner the ability to better understand the modes of dataveillance. This journey begins with being aware of the digital trail data that forms the key ingredient of this digital panopticon, leading to an understanding of data use and informed consent.

The audit of ad preferences on smartphone apps empowered participants to take control of their "data selves" (Lupton 2020) as they discovered tools to manage ad customization and to surveil Facebook and advertisers. The participants' audit can be used as a research tool to expose users to data and algorithmic literacy through an intuitive user interface. In the period following the fieldwork, other platforms like Google and YouTube have provided similar app-based interfaces for users to interact with data and affinities.

The self-administered audit of advertising preferences exhibits the potency of this method to expose participants to the entire array of algorithmic outputs. In a short span of a few minutes, participants were able to discover data ranging from their algorithmic interest profile, the

widespread residency of their personal information to a data trail of their advertising exposure and engagement record. During this research, Facebook was the only platform where the participants could encounter their “data selves” with ease. These account audits could be used to compare the accessibility of data and privacy control measures in a future cross-cultural and cross-platform research study.

Prior to the audit, some participants had expressed their worries over a lack of control and understanding about the data practices of digital platforms and felt a sense of losing their anonymity in the digital world. Some participants of this research had expectations that government should regulate data practices of digital platforms, and this was similar to findings of Steedman et al (2020). Ignorance about the revenue model of the digital platforms contributed to the anxieties of the participants when they encountered highly personalised advertising. These concerns were not universal and some of the participants felt their data was safe with large digital platforms that were more accountable as compared to government organizations.

Industry Experts working in media agencies and the marketing function of advertisers highlighted various sources of digital data that are used for creating affinity audience. They also described the state of privacy discussions with regards to online advertising. Industry participants were able to enumerate the complexity of advertising technology, algorithmic prediction, data brokers, the role of identifiers and tracking technologies to create a comprehensive profile of users. Amongst Everyday Users, there was a vernacular understanding that the only data available with the platforms was the one provided by them. Participants were not able to understand the profound implications of signing EULA and the level of interconnectedness between seemingly disparate parts of their digital traces.

The Industry Experts were asked about the level of discussion about consumer data privacy, and they clearly stated that these discussions were not common place in India. The industry participants believed that it would take some time for India to pass a comprehensive data privacy legislation like General Data Protection Regulation (GDPR). This was very evident in the interviews with industry experts who agreed that generally there were not many discussions about consumer data privacy in the media and advertising planning meetings. GDPR was something that happened in Europe and except for platforms owned by multinational corporations, training or discussions about privacy were rare. Given the global ownership of media agencies, the most popular digital media platforms and top advertisers, a sophisticated advertising delivery system is exposed to the population with a vernacular understanding of privacy, creating a fertile ground

for a privacy chasm. The use of technology for corporate surveillance or surveillance by the government did not surface in discussions.

7.4 Surveillance

Bentham's panopticon is used as the building block in literature about advertising as surveillance. Rewards and punishment in a panoptic prison induce desired behaviour change amongst the target population and is delivered through a perception of continuous observation. The observed are made aware of the continued presence of the observer but cannot monitor the existence or identity of the observer. Scholars (Gandy 1989, Dandeker 1990, Bogart 1996, Campbell and Carlson 2002, Elmer 2003, Mayer 2009, Saulles and Horner 2011) have labelled corporate data gathering practices as a modern-day panopticon. There are similarities as well as differences between the deployment of panoptic surveillance and advertising technology to achieve compliance to a desired behaviour.

There is commonality between a panopticon and advertising tech as, in both the cases the subject is under continuous observation. The two diverge in other aspects as, in the case of a panopticon, the mere perception of being under observation is expected to induce behaviour change, whereas in the case of ad tech, behaviour change is attempted by persuasion through advertising communication. In the case of Bentham’s panopticon, the act of observation is supposed to coerce the desired behaviour and is explicit. In the case of advertising technology, the observer attempts to remain invisible and deploys a visible stimulus to achieve the desired behaviour.

The act of selecting a suitable stimulant (advertisement) for a curated subset of the general population is where the surveillance is revealed to the observed. The identity of the communicator is revealed to the target audience through the stimulus. The process of curating a desired subset of the population is never revealed in an obvious manner. According to Berry (2011), the code that acts both as an actor and a vessel holding the data drives the invisible layer of advertising technology. This invisible layer of identifiers that connects an individual across the internet has been a focus of literary critique and regulators. Progressively, this scrutiny has resulted in hardware and software services increasing user visibility of permissions and control over preferences.

Scholars have compared personalised advertising to a consumer panopticon and the smartphone has been labelled as a mobile panopticon. Participants were generally appreciative of advertising that was relevant to their personal situation and aided their information search about an upcoming purchase decision. Some participants held a view that advertising technology should not know them personally and they liked the random advertising delivered via traditional

television. Other participants held a favourable view of personalised advertising when it provided information that aided in decision making and was supplemental to the primary activity. Advertising that hindered the primary experience like video, news and entertainment was disliked. However, highly entertaining advertising that is not repeated very often was acceptable even if it interrupted the main purpose of engaging with a digital platform. Television advertising was liked for its randomness, but participants felt a lack of control, whereas participants enjoyed the level of control offered by some digital advertising formats. In short, relevant advertising with control over the viewing experience was considered a desirable feature of the digital experience.

Advertising as surveillance is a focus area for academic research and algorithmic trading of an audience commodity is considered a dystopian functionality of profiling and sorting technologies. Analysis concluded that participants were not equipped to discover information to decode the underlying layer of advertising technology that funds most of the recognizable digital platforms such as Amazon, Facebook, Google, Instagram, TikTok and YouTube. In the absence of this information, people used heuristics to understand the ability of digital platforms to deliver advertising that is so relevant to their current personal situation.

Scholars have documented the common-sense explanations for complex phenomena like nano technology, journalism, and the curation of social media feeds. My research outlined a similar explanation for data gathering practices of digital platforms and called it “folk theory of data gathering”. A lack of data awareness and the inability to understand the workings of predictive advertising technology resulted in participants citing numerous examples where they were served advertising about topics that they had discussed in the vicinity of their smartphones. There was a widespread belief that smartphone apps such as Facebook and Instagram were listening to their conversations. Participants reached this conclusion by combining the repeated prompts by mobile apps to access cameras and microphones with advertising that was intimate to their personal circumstances.

Participants observed that topical advertising was served via digital platforms like Facebook, Google, and Instagram immediately after a discussion about the same topic. They theorised that for topical advertising to be served after a conversation about the topic in the proximity of a mobile phone, apps such as Facebook and Instagram must be listening to their conversations. This common-sense understanding about accuracy of customised advertising delivered by some of the largest digital platforms such as Google, Facebook and Instagram is called The Folk Theory of Customised Advertising. This is also referred to as audio snooping by some participants.

This research revealed other folk theories that highlighted the anxieties of the participants while dealing with complex information and the technology environment. These were about 1) the

prevalence of hackers who could easily steal their data as apps were seeking so many permissions, 2) TrueCaller uploading the contact lists of all the users into a centralised database to identify callers with scary accuracy and 3) Facebook wanting to add an Aadhar number to all the user accounts. All these are grouped under folk theories of data gathering.

Industry Experts were able to establish the process of creating personalised advertising, but without access to this level of data literacy, users are left with folk theories to understand these complex systems. Industry Experts described an extensive array of practices that were used to identify the right target audience for their campaign using algorithms running on large databases built on behavioural and transactional data. Industry participants also outlined methods used to identify cohorts through these databases to create an accurate profile of internet users enabling them to target people based on the affinity to a product category.

Participants described AFDP and services provided by them as essential to their day-to-day functioning. Participants outlined the portability of news and entertainment, location-based services, ride sharing services, geo-sensitive search results, social connections, and customised contextual advertising as the key benefits of mobile apps. There was a perceptible desire to have their data used for customised advertising and make the technology work in saving time, effort, and money. Relevance was an important factor in generating a positive attitude towards advertising.

Overall, favourability of advertising varied across platforms and was based on thematic and endemic suitability. Platforms such as Search that carried textual and contextual advertising were preferred, as they enhanced the purposive nature of engagement. While advertising on video entertainment platforms was viewed as a distraction, it was tolerated as a trade-off for not subscribing to ad-free premium services. Relevance, context, control over advertising experience, and compatibility with the native format cultivated positive attitudes towards advertising. Some participants spontaneously brought up their perceived invasion of privacy and introduced the concept of audio snooping into the discussion across multiple group interviews. While the participants seemed to enjoy personalised recommendations, customised advertising that was intimate to their personal situations caused great deal of anxiety. The anxieties about the uncanny accuracy of personalised advertising gave rise to apprehensions about audio, video, and locational surveillance. This disquiet manifests in common-sense folk theories about data gathering.

The research outlines a requirement for an easy-to-use toolkit for everyday smartphone users to become algorithm-aware, privacy conscious and develop an ability to provide informed consent. This research found a sense of resignation and a lack of control over their data usage and an

apathy towards informed consent. However, this research also highlighted the enabling nature of digital technologies for participants in India. At times participants were positively impacted by advertising technology and viewed personalised ads as information. They were able to utilize the power of ad serving algorithms to surveil large corporates and extract value by improving their online experiences. Some participants believed that personalised digital advertising was something that they were entitled to in return for sharing such detailed data about themselves.

7.5 Research Contributions and Implications

7.5.1 Summary of Research Findings

In this section I will summarize the main findings of this research.

FOMO, TINC, TINO and TINT: They are ways of describing how participants felt about their data practices of digital platforms like Amazon, Facebook, Google, and Instagram. These are also mitigation strategies as described in Table 2. These answer the RQ1 about the data experiences of young Indian smartphone users.

Table 2 describes key tactics deployed by participants when faced with a litany of factors related to their datafication. There was an understanding amongst participants that digital platforms are collecting data about their usage, but there was a sense of apathy and resignation when it came to making efforts to understand their “data doubles”. Participants mitigated their data anxieties by resorting to tactics like TINO, FOMO, TINT.

TINO is deployed when an app like Uber is downloaded to the mobile device when a ride is needed. As they felt that there is no alternative to the ride sharing service, accepting all the terms and conditions without reading is considered acceptable. Similar is the case with other apps listed in the table above. TINO is also used to justify using apps such as Facebook and Instagram even when some participants felt that these apps were serving advertisements related to their conversation topics.

FOMO is deployed in in the context of social as well as e-commerce apps. Participants felt that if they move away from social media apps like Instagram or Facebook they will miss out on social conversations. Instagram straddled across social connections and connecting with brands. Participants recalled that they received information about their favourite brands and special offers from this platform and moving away from it would mean missing out on these announcements.

TINT is a derivative of TINO and FOMO. As participants imagined that there is no alternative to certain apps while few other apps are installed to avoid FOMO. Thus, the total number of apps on a smartphone balloon up. Such a large number of apps installed on their phones, make it difficult for the participants to actively manage their privacy settings. A common theme that runs across all the groups is that there is no time to actively manage so many apps and keep track of their changing usage policies.

TINC is similar to TINT and it is used to alleviate issues around lack of active participation in digital trail data management. It was felt that if they do not have any control over their digital usage data, why waste efforts in trying to manage it. Some participants noticed that the ads followed them across the internet as they have no control over their own data. There was a sense of inevitability, if they manage their privacy on one app, the advertisers will be able to reach them via another application. Inevitability often led to apathy.

Self-audit ad preferences on Facebook: This is a powerful methodological tool that surfaces multiple aspects of algorithmic sort to participants. By browsing through their privacy controls and ad preferences, participants were able to 1) examine the interest cohorts that they were placed in by the digital platform, 2) grasp the availability of their personal information with advertisers, 3) develop awareness about the controls available to manage personalization of advertising, 4) audit their past interactions with advertising on the platform and 5) reflect on the digital data trails that they created while using a digital platform.

Figure 2 shows one such path from settings and privacy menu to a list of advertisers that possess the data about the individual or have used off Facebook activity to target them on Facebook mobile app. This proved to be a powerful tool to understand the first impressions of participants while they were being exposed to functioning of personalised targeting and advertising. This experience generated a wide range of reactions ranging from, it’s OK if they have all this data to why would they share so much detail with us. Some participants perceived this as a learning experience and thanked the researcher, as they learned something new from this interaction.

This method of self-audit and reporting, provides an innovative method to explore data experiences of participants as compared to intrusive app installation of Pybus et al (2015), provides an alternative for understanding data practices of Facebook (Kennedy 2016, pp. 172–173), address the challenges of researching Facebook raised by Varis (2016, p. 63) about "a full ethnography of such a multi-channelled site", provides an alternative approach for conducting "stakeholder ethnography" (Pink et al 2017, p.176) and would have simplified participants recruitment and new software development for the research of Skeggs and Yuill (2016 b). Shklovski et al (2014) conducted two research studies to understand the experience of users in

western Europe when confronted with data collection practices of mobile apps. This method of self-audit of ad preferences on apps already installed and used by participants is an improvement over lab studies.

Determining trust threshold: Participants used multiple factors to determine if a mobile application should be trusted. Key factors that were considered by the participants of my research: 1) tenure of existence, 2) brand or corporate entity owning an app, 3) usage of the app within their cohort, 4) reviews on app store, 5) acceptable level of permissions sought by the mobile app and 6) absence of a corrosive experience with the app. Some participants cited FOMO on social interactions and discount offers as a reason for pushing certain apps over the trust threshold. Passing the trust threshold was important for the participants as they were comfortable with crossing privacy redlines for apps that they trusted. Privacy redlines were 1) sharing their financial details like bank accounts and credit card details and 2) permission to access contact list, emails, location, and photos.

Folk Theories of Data Gathering: It was observed that participants were not able to understand the data gathering practices of digital platforms and underlying layer of advertising technology that connected their identities across the internet. Faced with a complex technology ecosystem, participants used their common-sense understanding to develop folk theories of data gathering. Folk theory of customised advertising was discussed across multiple focus groups in Mumbai and Delhi. Participants believed that they were receiving advertising on digital platforms like Amazon, Facebook, Google Search, and Instagram about topics that they had recently discussed in the vicinity of their mobile phones. Their theorised that the only possible answer was that these applications were listening to their conversations via their mobile phones and serving topical advertisements. Participants called it audio snooping. There was no evidence of this practice (Facebook and Google listening to private conversations for personalization of advertising) from abovementioned platforms in the interviews of Industry Experts, academic literature or from any privacy regulator.

Data Experiences in Indian Context: My research found some distinctive insights about data experiences of Everyday Users of digital platforms in India.

1) Trust in Digital Platforms: Participants exhibited higher level of trust in digital platforms to keep their data safe as compared to the government. This trust was earned when the apps passed the trust threshold as described above. Perceived accountability of privately owned corporations was believed to be the main reason for participants to entrust their data to digital platforms. Participants used words such as naïve and clueless to describe the government.

2) Mobile phone is not a personal device: Mobile phone is a social device in the participants' household. Participants worried that predictive algorithms could reveal highly personal information through recommendations and personalised advertising. Some female participants were worried that their search history could reveal embarrassing topics to other members of their household. Similar concerns were expressed by other participants about their YouTube watch history. Overall, participants were worried that personalised advertising or content recommendations could make undesired exposure to other family members as anyone in the household could pick up a phone to use it.
3) Short history of datafication: India has a very short history of electronic payment systems or of a credit bureau. In section 4.3 participants compared the anonymous offline shopping to online shopping where they had to provide a lot of sensitive information to get home delivery of goods purchased online. They also surface an important point about the short history of online shopping itself. This newfound convenience, short history of datafication, little regulatory oversight and an absence of debates about datafication are largely absent from public discourse. This leads to very low level of sensitivity towards hazards of sharing their PIIs. Some participants felt that their phone number was not a piece of sensitive information and they freely shared it across the internet to get free trials and discounts offers. A lack of data activism can also be attributed to abovesaid factors.
4) Users demand personalised advertising: Throughout this research, the participants expressed an affirmative approach towards customised advertising. Participants expected to be served with customised advertising as they found it relevant for their purposive activities like surveiling the e-commerce website for discounts and information about new launches from preferred brands.
5) Lack of data protection laws: India does not have a data protection law, and this reflects political priorities with regards to user data. Despite a lack of legislation or regulatory oversights, Indian users of global digital platforms like Amazon, Facebook, Google, and Instagram benefitted from changes in global privacy landscape. Self-audit of mobile applications was only possible because Facebook app reflected their global move towards providing users control over their ad experiences. Participants acknowledged the trend of increased control over permissions to use hardware features like microphones and cameras. In the past, participants were not even aware of the fact that these permissions were hidden in the terms and conditions that they had accepted. They appreciated increasing transparency over the use of hardware features by digital platforms but did not understand the context in which this was happening. Absence of public discussion in India about digital data privacy meant that the context of greater control over their digital data trails was unclear. The lack of context was evident when participants did not have a framework to process a sudden

discovery of detailed information during self-audit of Facebook. Some of the participants were amused by the fact that Facebook provided so much transparency and others felt that Facebook was washing its hands off all responsibility by placing the burden of managing their privacy on the users. This exercise highlighted the changing privacy landscape that the participants had experienced over the years and how Indian users were benefiting from pressures elsewhere in the world.

7.5.2 Implications

The output as discussed in the previous section has implications across policy, industry practice, academia, and research methods. Several policy and regulatory initiatives across geographies are aiming to provide better controls over data privacy and data experiences to the users of digital platforms. A Parliamentary Committee had been constituted to examine The Personal Data Protection Bill that was presented in the Indian parliament on 11th December 2019 (Indian Parliament, 2019) and this research could contribute to the understanding of urban Indian audiences and their expectations from regulators.

As indicated by scholars (Kennedy 2018), the data activism that drives the legislative agenda is at times driven by data elites theorising about everyday data experiences of common people and this research brings out the affirmative nature of such technologies and de-Westernizes the narrative of exploitation. This research contributes to similar studies carried out in Australia, the United Kingdom, the United States, and other European countries. This research makes some important contributions to the understanding of Indian audiences and their anxieties over the use of their data and hardware permissions.

There is an urgent requirement for digital platforms to proactively guide their users through in app navigation so that the process of privacy management is easy to understand and access. This would enable the users to improve their understanding about the working of advertising technology and enhance their data experiences. In the absence of endogenous source of information, users are motivated to look for answers through exogenous sources and develop folk theories by equating chronology with causality. The widespread prevalence of folk theory of data gathering is one such example that fuels the techno-dystopian narrative. These sentiments have been reported in mainline newspapers as well as documentaries exploring contemporary data experiences of everyday users. The Great Hack (2019) shows at 2:19 minutes that majority of the students of Professor David Carroll felt that their phones were listening to them. The Wall Street Journal (2021) listed numerous TikTok users describing the same feeling.

Key concepts that emerged from the analysis were:

- Folk theories of data gathering
- The role of relevance and control of experience in perception of digital advertising
- The concept of a privacy divide driven by the cost of technology and the regulatory environment in India
- The concept of an inverted advertising panopticon
- Users' rationale for avoiding knowledge about terms and conditions in EULA
- A methodological approach of user audits of advertising preferences.

7.6 Limitations

The purposive nature of sampling means that the results of this research are not representative of the entire population of 18 to 34 year old people in Mumbai and Delhi. Participants recruitment was conducted across different areas of Delhi and Mumbai, but the research does not represent all the social classes residing in these two cities. As qualitative research methods were used for data gathering, they could be biased by the snowballing nature of recruitment as well as the confirmation bias that occurs in a group interview situation. The researcher who conducted the focus group discussions took all precautions to ensure that the conversation was not dominated by a particular individual/participant; however, it was not possible to be successful all the time. The concept of data was varied amongst the participants in a group and for a common understanding was contingent on the composition of the group. There were some participants who thought of internet data plans whilst describing their experience with data and it took some time to understand the complex three-way exchange between themselves, the digital platforms, and the advertisers.

7.7 Areas of Future Research

The design of this research, where the researcher interviewed Industry Experts who could shed light on industry practices that create this level of customization, helped in understanding the nuances of technologies and how myths spread through society as folk theories were found in different areas of the same city but also in different cities. This design acted as a mechanism providing evidence for the spread of theories across wider society.

- The role of hardware in enabling surveillance of users: Participants discussed in detail various permissions sought by mobile applications. Scholars have written about the increasing datafication of audiences which has been made possible due to the explosion in metadata capture. Despite massive advancements in the capabilities of mobile hardware to accurately predict geo location via GPS signals, there is scant research

available on the enhanced capabilities of hardware devices and the ability of mobile apps to access this metadata via operating systems.

- The role of operating systems and walled gardens around app stores resulting in an increased surveillance of users: Operating systems have evolved and along with the hardware gathered copious amount of data about users to provide better personalised services. Discussions with industry experts revealed that there are no industry standards about the amount of data that a single mobile application can access from the pool of data that is gathered by the operating system.
- A detailed study of processes in advertising technology to understand the interactions between various sources of consumer data. According to Industry Experts there are three sources of data, first-party, second-party and third-party and ad tech aggregates data from all these sources to create detailed profiles of consumers. Some of the Everyday Users were unaware of this underlaying layer of ad tech and its ability to generate personalised advertising. There is a possibility for future research to uncover the relationships between various elements of advertising technology with the long-time culprit, the cookie. The cookie has been gradually deprecated by Safari and Firefox browsers. Google's Chrome has announced the sunset date for the third-party cookie to be 2023. There is a requirement to understand the new data ecosystem that is evolving due to these challenges to the legacy systems. There is scant research into Apple's Intelligent Tracking Prevention initiative that was launched in 2017 and this highlights the lag in practice and academia.
- A quantitative study to generalise the folk theory of data gathering: This is one of the most significant contributions of the research and a quantitative study to generalise the findings is an area for further research.
- An autoethnography of a frequent user of mobile applications to understand the journey of a data user in the rapidly changing technology environment. This combined with an exploration of the impact of regulation on the experiences of data privacy and controls associated with it will be an important area to explore.

Appendix A Discussion Guide

ERGO number: 52140

A.1 Focus Group Discussions

Questions for Everyday Users in Delhi and Mumbai

1. What all forms of media do you consume in your day to day life?

2. What are the most favourite CPG brands?

3. What applications and websites do you consume everyday and what is the gratification from them?

4. Is advertising more acceptable in one format over other and why?

5. What do you understand about privacy with regards to online space like social media networks, news sites, video entertainment, navigation or online search?

6. Are you aware about what all data is gathered by digital media formats?

7. How do they feel about the safety and security of the data?

8. How do you think all the online platforms make money, what is their business model?

9. How do they feel about being bought and sold as commodities?

10. Explore how useful digital platforms are to their personal, work or social life.

11. Do they feel exploited by the digital media platforms? Or they feel that they are useful?

12. Probe about use of digital platforms to make purchase decisions.

13. Ask if they have ever tried to delete their personal information on a digital platform.

14. Ask them to describe some multi media advertising campaigns.

15. Is there a difference between memorability of digital or traditional media formats?

16. How do they rank the control of their advertising experience? E.g. Skipping ads.

17. Do they feel if the ads are watching them and how do they feel about it?

18. What all data they are comfortable if the platforms are sharing with advertisers or other third parties and where is their red line?

19. Ask the respondents to do a live exercise to access an entertainment site/ app, a news site/ app, a social media site or app and online shopping site or app and a search site or app and observe their behaviour.

Additional Questions for in-Depth Interviews

1. Are they aware of where they can find the data about their online usage?

2. Understand their feeling about the exchange of their personal usage data with advertisers for free or subsidised access to different digital platforms

3. Guide them towards the data access menus in their top three online applications.

a. Take a note of all the data that is available and depth and granularity available.

b. Probe again about their feelings towards digital platforms and how they feel now about buying and selling of audience data with advertisers and other third parties.

A.2 In-Depth Interview Industry Experts

1. Are they satisfied with the amount of information available for media planning, buying and selling?

2. How has this practice changed over the last two decades and what was the inflection point?

3. How do the media planning and measurement datasets for traditional and digital media differ?

4. What is the consumer usage data that is gathered and available at an individual level and aggregated level for media agencies, advertisers and digital platforms?

5. How important is informed consent in collection of consumer data?

6. What are the processes and practices in place for collecting and managing user data?

7. How is data collected and assigned to various stages of consumer decision journey?

8. How important is quantitative versus qualitative data about the consumers?

9. What is the process of ingesting this data into media planning and measurement tools?

10. How has the privacy debate taking shape amongst practitioners?

11. Are these discussions institutionalised?

12. What are the top ten most important data fields that are crucial for a successful media planning and buying practice?

13. Do these differ for different product categories?

14. What typical challenges do CPG advertisers face in the increasingly digital world as most of their sales are still offline?

Appendix B Code Book

Code	Description	Files	References
Ad Fraud	Discussions about ad fraud were coded here	1	1
Audiences	How digital audiences are being curated as per the consumers.	6	7
Consumer		12	121
Advertising Attitudes	What do the consumers feel about advertising. Main focus is digital advertising.	11	29
Consumer Concerns			
Consent	Do the consumers feel that they are asked for consent before their data is used or collected.	3	4
Exploitation	Do the consumers feel exploited by the digital platforms.	8	19
Dataveillance	Do the consumers feel under surveillance for their data or using their data.	4	11
Regulation	Who should be regulating digital platforms with regards to privacy.	4	5
Sharing	Concerns about sharing their data with digital platforms.	6	12
Consumer Data		12	92
Access	How much consumers have access to their usage data on apps and sites.	3	5
Awareness	Awareness of consumers about the data being collected by apps and sites.	11	29
Control	How much control do customers have on their data.	9	23
Exchange	Discussion around the exchange of consumer data for free access to apps.	8	23
Safety	What do consumers feel about the safety of their data on internet and with specific apps.	5	12
Consumer Privacy		0	0
Privacy Concerns	Key concerns with regards to privacy,	10	20
Privacy Red line	Data privacy red lines, which personal data consumers do not want digital platforms to have.	8	9
Consumer Brands	CPG brands that consumers use. Warm up discussion.	6	8
Consumer Business Model	Business model of free to use digital platforms.	11	16

Code	Description	Files	References
Consumer Digital Business Model	Consumer understanding of business model of free online apps.	4	4
Consumer Mobile Usage	General mobile usage and warm up discussion.	12	22
Consumer Myths	Myths surrounding the use of internet and apps with regards to data and privacy.	2	7
Consumer Online Shopping	Discussions about online shopping experience.	4	6
Consumer Recommendation	What do consumers feel about the recommendations that they get from various digital platforms.	0	0
Consumer Trust	Discussions about trust of consumers in digital platforms or internet in general.	5	9
Customer Journey	Discussions about consumer purchase journey.	0	0
Customer Journey Auto	Customer purchase journey for Auto Industry.	1	2
Customer Journey Beverages	Customer journey for food and beverages industry.	1	1
Customer Journey Insurance	Consumer purchase journey for Insurance.	1	2
Data Availability	Data availability and its breakdown is coded here.	3	5
Cross Platform Data	What is the type of cross platform data is available with the marketer or agency.	2	4
Data availability agency	Discussion around what data is available with the media agency.	1	2
Data Process	Discussion about who owns the process of aggregating consumer data for use in media planning.	3	4
General	General discussions about data availability.	1	1
Programatic	Discussions about programatic media planning and buying.	1	1
Qualitative Data	Use of qualitative data in digital media planning.	2	2
Data availability platform	Discussions about availability of consumer data with digital platforms is coded here.	1	1
Data Ecosystem	Discussions about the digital media ecosystem are coded here.	2	4
Data Usage	All discussions related to usage of consumer data by communication	1	1

Code	Description	Files	References
	industry are coded here.		
Crucial Data Sets	Most important data for the media planning.	5	9
Data Integration Marketer	Discussions about how data from various sources is integrated for planning purposes.	2	3
Data Usage Beverages	Usage of consumer data in Food and Beverage industry is coded here.	1	4
Data Usage Insurance	Use of consumer data in the Insurance industry is coded here.	1	1
Data Usage Process	Discussions about the process of collecting and using consumer's data are coded here.	2	2
Remarketing	Discussions about remarking practices are coded here.	1	1
Remarketing and Regulation	Discussions about remarking and practices/ regulations are coded here.	1	1
FB Audit	Discussions related to Facebook audit are coded here.	1	1
FB Audit DI	Discussions about Facebook audit during depth interviews are coded here.	1	1
FB Audit FGD	Discussions about Facebook audit during focus group discussions are coded here.	5	10
FB Audit Reaction	Consumer reactions to Facebook audit are coded here.	2	2
History	Views of Industry experts on how data availability has changed over the past decade.	4	4
Measurement	Views of professionals on industry practices on audience measurement.	2	3
Media Planning	Views of Professionals on Media Planning.	2	5
Observations consumer	Observations of the researchers during his interaction with consumers.	1	1
Observations Consumer FGD	Observations during Focus Group Discussions.	6	61
Observations Consumers DI	Depth Interview observations.	5	16
Observations Professionals	Observations of researcher on the discussions with professionals.	3	5
Digital Echo	Professionals describing how the consumer data is captured in the digital world.	1	2

Code	Description	Files	References
Privacy	Discussions that professionals have about privacy in their organizations.	6	11
Qualitative Data	Views of professionals about use of qualitative data in media planning and buying.	3	5
ROPO	Research Online Purchase Offline	1	1
Trust	What does professionals do to improve trust of consumers.	2	3
TV Advertising	What do respondents feel about TV Advertising	4	5

Appendix C Participant Classification

Assigned Name	Age Group	City	Gender	Industry Participant
Group 1\G1P1	18 to 24	Delhi	Male	Yes
Group 1\G1P2	18 to 24	Delhi	Male	No
Group 1\G1P3	18 to 24	Delhi	Male	Yes
Group 1\G1P4	18 to 24	Delhi	Male	Yes
Group 1\G1P5	18 to 24	Delhi	Male	Yes
Group 1\G1P6	18 to 24	Delhi	Male	Yes
Group 2\G2P1	18 to 24	Delhi	Female	No
Group 2\G2P2	18 to 24	Delhi	Female	No
Group 2\G2P3	18 to 24	Delhi	Female	No
Group 2\G2P4	18 to 24	Delhi	Female	No
Group 2\G2P5	18 to 24	Delhi	Female	No
Group 2\G2P6	18 to 24	Delhi	Female	No
Group 3\G3P1	25 to 34	Delhi	Female	No
Group 3\G3P2	25 to 34	Delhi	Female	No
Group 3\G3P3	25 to 34	Delhi	Female	No
Group 3\G3P4	25 to 34	Delhi	Female	No
Group 3\G3P5	25 to 34	Delhi	Female	No
Group 3\G3P6	25 to 34	Delhi	Female	No
Group 4\G4P1	25 to 34	Delhi	Male	No
Group 4\G4P2	25 to 34	Delhi	Male	No
Group 4\G4P3	25 to 34	Delhi	Male	No
Group 4\G4P4	25 to 34	Delhi	Male	No
Group 4\G4P5	25 to 34	Delhi	Male	No
Group 4\G4P6	25 to 34	Delhi	Male	No
Group 5\G5P1	18 to 24	Mumbai	Female	Yes
Group 5\G5P2	18 to 24	Mumbai	Female	Yes
Group 5\G5P3	25 to 34	Mumbai	Female	No
Group 5\G5P4	18 to 24	Mumbai	Female	Yes
Group 5\G5P5	18 to 24	Mumbai	Female	Yes
Group 6\G6P1	25 to 34	Mumbai	Female	No
Group 6\G6P2	25 to 34	Mumbai	Female	No

Assigned Name	Age Group	City	Gender	Industry Participant
Group 6\G6P3	25 to 34	Mumbai	Female	No
Group 6\G6P4	25 to 34	Mumbai	Female	No
Group 6\G6P5	25 to 34	Mumbai	Female	No
Group 6\G6P6	25 to 34	Mumbai	Female	No
Group 7\G7P1	25 to 34	Mumbai	Male	No
Group 7\G7P2	25 to 34	Mumbai	Male	No
Group 7\G7P3	25 to 34	Mumbai	Male	No
Group 7\G7P4	25 to 34	Mumbai	Male	No
Group 7\G7P5	25 to 34	Mumbai	Male	No
Group 7\G7P6	25 to 34	Mumbai	Male	No
In-Depth Interviews\DI F 27	25 to 34	Delhi	Female	Yes
In-Depth Interviews\DI F23	18 to 24	Delhi	Female	Yes
In-Depth Interviews\DI F32	25 to 34	Mumbai	Female	Yes
In-Depth Interviews\DI M 25	25 to 34	Delhi	Male	Yes
In-Depth Interviews\DI M 27	25 to 34	Delhi	Male	Yes

Appendix D Ethical Documents

D.1 Participant Information Sheet (PIS) Depth Interview

Study Title: An Exploration of Processes Used by Consumer Packaged Goods Marketers and Digital Platforms to Create Custom Audiences in the Digital World and its Implications for Advertising

Researcher: Ravi Dixit

ERGO number: 52140

You are being invited to take part in the above research study. To help you decide whether you would like to take part or not, it is important that you understand why the research is being done and what it will involve. Please read the information below carefully and ask questions if anything is not clear or you would like more information before you decide to take part in this research. You may like to discuss it with others but it is up to you to decide whether or not to take part. If you are happy to participate you will be asked to sign a consent form.

Q1. What is the research about?

This is a student project for my PhD and I will be asking questions about your usage of internet for day to day task like using it for research, entertainment, accessing social media. I will also ask you about how you feel about providing your personal usage data with various digital platforms and the reason why you feel comfortable or otherwise in doing so. The information that you provide will be used to understand the exchange of free internet services for personal usage data.

Q2. Why have I been asked to participate?

There will be a total of 48 participants in this study in the age range of 18 to 34 from Mumbai and Delhi. As you are in the target age group you are being asked to be a participant in this study.

Q3. What will happen to me if I take part?

You will be a central location for an interview with the researcher and you will express your opinion about the topics stated above. There will be an audio recording of the discussion so that it can be transcribed so that all the information is captured accurately.

Total time commitment that will be required of you is 25 minutes.

Q4. Are there any benefits in my taking part?

Your participation will help me with my study and it will further the academic knowledge about the usage of internet services.

Q.5 Are there any risks involved?

There are not risks involved in participating in this study.

Q.6 What data will be collected?

Your personal data about age, gender and socio economic class will be collected along with your views about the topic under discussion. All the data will be stored on secure university systems.

Q.7 Will my participation be confidential?

Your participation and the information we collect about you during the course of the research will be kept strictly confidential.

Only members of the research team and responsible members of the University of Southampton may be given access to data about you for monitoring purposes and/or to carry out an audit of the study to ensure that the research is complying with applicable regulations. Individuals from regulatory authorities (people who check that we are carrying out the study correctly) may require access to your data. All of these people have a duty to keep your information, as a research participant, strictly confidential.

All The audio recordings will be destroyed after they have been transcribed by the researcher and consent forms will be stored in locked cabinets at the university.

Q.8 Do I have to take part?

No, it is entirely up to you to decide whether or not to take part. If you decide you want to take part, you will need to sign a consent form to show you have agreed to take part. You will have to fill in the consent form now if you are willing to participate.

Q.9 What happens if I change my mind?

You have the right to change your mind and withdraw at any time without giving a reason and without your participant rights being affected. If you do not fill in the form we will not have any data from you. The information collected about you up to this point may still be used for the purposes of achieving the objectives of the study only

Q.10 What will happen to the results of the research?

Your personal details will remain strictly confidential. Research findings made available in any reports or publications will not include information that can directly identify you without your specific consent.

Q.11 Where can I get more information?

You can reach the researcher at rd7g15@soton.ac.uk for any information at any time. If you want a copy of the final report please write to the researcher.

Q. 12 What happens if there is a problem?

If you have a concern about any aspect of this study, you should speak to the researchers who will do their best to answer your questions.

If you remain unhappy or have a complaint about any aspect of this study, please contact the University of Southampton Research Integrity and Governance Manager (023 8059 5058, rgoinfo@soton.ac.uk).

You can reach the researcher at rd7g15@soton.ac.uk for any information at any time.

Data Protection Privacy Notice

The University of Southampton conducts research to the highest standards of research integrity. As a publicly-funded organisation, the University has to ensure that it is in the public interest when we use personally-identifiable information about people who have agreed to take part in research. This means that when you agree to take part in a research study, we will use information about you in the ways needed, and for the purposes specified, to conduct and complete the research project. Under data protection law, 'Personal data' means any information that relates to and is capable of identifying a living individual. The University's data protection policy governing the use of personal data by the University can be found on its website (https://www.southampton.ac.uk/legalservices/what-we-do/data-protection-and-foi.page).

This Participant Information Sheet tells you what data will be collected for this project and whether this includes any personal data. Please ask the research team if you have any questions or are unclear what data is being collected about you.

Our privacy notice for research participants provides more information on how the University of Southampton collects and uses your personal data when you take part in one of our research projects and can be found at http://www.southampton.ac.uk/assets/sharepoint/intranet/ls/Public/Research%20and%20Integrity%20Privacy%20Notice/Privacy%20Notice%20for%20Research%20Participants.pdf

Any personal data we collect in this study will be used only for the purposes of carrying out our research and will be handled according to the University's policies in line with data protection law. If any personal data is used from which you can be identified directly, it will not be disclosed to anyone else without your consent unless the University of Southampton is required by law to disclose it.

Data protection law requires us to have a valid legal reason ('lawful basis') to process and use your Personal data. The lawful basis for processing personal information in this research study is for the performance of a task carried out in the public interest. Personal data collected for research will not be used for any other purpose.

For the purposes of data protection law, the University of Southampton is the 'Data Controller' for this study, which means that we are responsible for looking after your information and using it properly. The University of Southampton will keep identifiable information about you for 10 years after the study has finished after which time any link between you and your information will be removed.

Thank you.

D.2 Participant Information Sheet (PIS) Focus Group Discussion

Study Title: An Exploration of Processes Used by Consumer Packaged Goods Marketers and Digital Platforms to Create Custom Audiences in the Digital World and its Implications for Advertising

Researcher: Ravi Dixit

ERGO number: 52140

You are being invited to take part in the above research study. To help you decide whether you would like to take part or not, it is important that you understand why the research is being done and what it will involve. Please read the information below carefully and ask questions if anything is not clear or you would like more information before you decide to take part in this research. You may like to discuss it with others but it is up to you to decide whether or not to take part. If you are happy to participate you will be asked to sign a consent form.

Q1. What is the research about?

This is a student project for my PhD and I will be asking questions about your usage of internet for day to day task like using it for research, entertainment, accessing social media. I will also ask you about how you feel about providing your personal usage data with various digital platforms and

the reason why you feel comfortable or otherwise in doing so. The information that you provide will be used to understand the exchange of free internet services for personal usage data.

Q2. Why have I been asked to participate?

There will be a total of 48 participants in this study in the age range of 18 to 34 from Mumbai and Delhi. As you are in the target age group you are being asked to be a participant in this study.

Q3. What will happen to me if I take part?

You will be a part of a group of up to 8 participants invited to a central location for a Group Discussion and you will express your opinion about the topics stated above. There will be an audio recording of the discussion so that it can be transcribed so that all the information is captured accurately.

Total time commitment that will be required of you is 120 minutes.

Q4. Are there any benefits in my taking part?

Your participation will help me with my study and it will further the academic knowledge about the usage of internet services.

Q5. Are there any risks involved?

There are not risks involved in participating in this study.

Q6. What data will be collected?

Your personal data about age, gender and socio economic class will be collected along with your views about the topic under discussion. All the data will be stored on secure university systems.

Q7. Will my participation be confidential?

Your participation and the information we collect about you during the course of the research will be kept strictly confidential.

Only members of the research team and responsible members of the University of Southampton may be given access to data about you for monitoring purposes and/or to carry out an audit of the study to ensure that the research is complying with applicable regulations. Individuals from regulatory authorities (people who check that we are carrying out the study correctly) may require access to your data. All of these people have a duty to keep your information, as a research participant, strictly confidential.

All The audio recordings will be destroyed after they have been transcribed by the researcher and consent forms will be stored in locked cabinets at the university.

Q8. Do I have to take part?

No, it is entirely up to you to decide whether or not to take part. If you decide you want to take part, you will need to sign a consent form to show you have agreed to take part. You will have to fill in the consent form now if you are willing to participate.

Q9. What happens if I change my mind?

You have the right to change your mind and withdraw at any time without giving a reason and without your participant rights being affected. If you do not fill in the form we will not have any data from you. The information collected about you up to this point may still be used for the purposes of achieving the objectives of the study only

Q10. What will happen to the results of the research?

Your personal details will remain strictly confidential. Research findings made available in any reports or publications will not include information that can directly identify you without your specific consent.

Q.11 Where can I get more information?

You can reach the researcher at rd7g15@soton.ac.uk for any information at any time.

Q12. What happens if there is a problem?

If you have a concern about any aspect of this study, you should speak to the researchers who will do their best to answer your questions.

If you remain unhappy or have a complaint about any aspect of this study, please contact the University of Southampton Research Integrity and Governance Manager (023 8059 5058, rgoinfo@soton.ac.uk).

You can reach the researcher at rd7g15@soton.ac.uk for any information at any time.

Data Protection Privacy Notice

The University of Southampton conducts research to the highest standards of research integrity. As a publicly-funded organisation, the University has to ensure that it is in the public interest when we use personally-identifiable information about people who have agreed to take part in research. This means that when you agree to take part in a research study, we will use

information about you in the ways needed, and for the purposes specified, to conduct and complete the research project. Under data protection law, 'Personal data' means any information that relates to and is capable of identifying a living individual. The University's data protection policy governing the use of personal data by the University can be found on its website (https://www.southampton.ac.uk/legalservices/what-we-do/data-protection-and-foi.page).

This Participant Information Sheet tells you what data will be collected for this project and whether this includes any personal data. Please ask the research team if you have any questions or are unclear what data is being collected about you.

Our privacy notice for research participants provides more information on how the University of Southampton collects and uses your personal data when you take part in one of our research projects and can be found at http://www.southampton.ac.uk/assets/sharepoint/intranet/ls/Public/Research%20and%20Integrity%20Privacy%20Notice/Privacy%20Notice%20for%20Research%20Participants.pdf

Any personal data we collect in this study will be used only for the purposes of carrying out our research and will be handled according to the University's policies in line with data protection law. If any personal data is used from which you can be identified directly, it will not be disclosed to anyone else without your consent unless the University of Southampton is required by law to disclose it.

Data protection law requires us to have a valid legal reason ('lawful basis') to process and use your Personal data. The lawful basis for processing personal information in this research study is for the performance of a task carried out in the public interest. Personal data collected for research will not be used for any other purpose.

For the purposes of data protection law, the University of Southampton is the 'Data Controller' for this study, which means that we are responsible for looking after your information and using it properly. The University of Southampton will keep identifiable information about you for 10 years after the study has finished after which time any link between you and your information will be removed.

Thank you.

D.3 Consent Form

Study title: An Exploration of Processes Used by Consumer Packaged Goods Marketers and Digital Platforms to Create Custom Audiences in the Digital World and its Implications for Advertising

Appendix D

Researcher name: Ravi Dixit

ERGO number: 52140

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet PIS V1 29072019 and have had the opportunity to ask questions about the study.	
I agree to take part in this research project and agree for my data to be used for the purpose of this study. I understand that these will be recorded using audio and written notes.	
I understand my participation is voluntary and I may withdraw at any time for any reason without my participation rights being affected.	
I understand that I may be quoted directly in reports of the research but that I will not be directly identified (e.g. that my name will not be used).	
I understand that I must keep the discussions confidential.	
I understand that my personal information collected about me such as my name or where I live will not be shared beyond the study team.	
I understand that should I withdraw from the study then the information collected about me up to this point may still be used for the purposes of achieving the objectives of the study only.	

Name of participant (print name)……

Signature of participant……

Date……. ……..

Name of researcher (print name)……

Signature of researcher ……

Date…….

List of References

Agrawal, R. (2018) India connected: how the smartphone is transforming the world's largest democracy. New York: Oxford University Press.

Anderson, C. (2009) 'Essay: the economics of giving it away', *The Wall Street Journal* (English edition)*,* 31 January [online]. Available at: https://www.wsj.com/articles/SB123335678420235003 (Accessed: 26 August 2020)

Andrejevic, M. (2007) *iSpy: surveillance and power in the interactive era.* Lawrence: University Press of Kansas.

Andrejevic, M. (2013) *Infoglut: how too much information is changing the way we think and know.* New York: Routledge.

Andrejevic, M. (2014) 'Big data, big questions| the big data divide', *International Journal of Communication*, 8, p.17.

Andrejevic, M., Hearn, A. and Kennedy, H. (2015) 'Cultural studies of data mining: introduction', *European Journal of Cultural Studies*, 18(4-5) pp. 379–394.

Arora, P. (2016) 'Bottom of the data pyramid: big data and the global south', *International Journal of Communication,* 10, p.19.

Athique, A. (2012) *Indian media*. Cambridge: Polity.

Auletta, K. (2018) *Frenemies: the epic disruption of the ad business (and everything else)*. London: Penguin Books.

Batra, A. (ed.) (2003) *Marketing whitebook 2003–2004.* New Delhi: BWBooks.

Batra, A. (ed.) (2017) *Marketing whitebook 2017–2018.* New Delhi: BWBooks.

Batra, R. and Ray, M.L. (1985) 'How advertising works at contact', *Psychological processes and advertising effects*, pp. 13–44.

Batra, R., Ahuvia, A. and Bagozzi, R.P. (2012) 'Brand love', *Journal of Marketing*, 76(2), pp. 1–16.

Bawden, D. (2008) 'Origins and concepts of digital literacy' in Lankshear, C. and Knobel, M. (eds.) *Digital literacies: concepts, policies and practices*, 30(2008), New York: Peter Lang, pp. 17–32.

List of References

Beer, D. (2009) 'Power through the algorithm? Participatory web cultures and the technological unconscious', N*ew Media & Society*, 11(6), pp. 985–1002.

Berry, D. (2011) *The philosophy of software: code and mediation in the digital age*. *E-book library* [online]. Available at: http://lib.myilibrary.com/?id=306769 (Accessed 14 March 2021)

Bodle, R. (2016) 'A critical theory of advertising as surveillance' in Hamilton, J. F., Bodle, R. and Korin, E. (eds.) *Explorations in critical studies of advertising*. New York: Routledge, pp. 138–152.

Bogard, W. (1996) *The simulation of surveillance: hyper-control in telematic societies*. Cambridge: Cambridge University Press.

Bolin, G. and Andersson Schwarz, J. (2015) 'Heuristics of the algorithm: big data, user interpretation and institutional translation', *Big Data & Society*, 2(2), pp. 1–12.

Borges, J.L. (1964) *Labyrinths: selected stories & other writings.* New York: New Directions Publishing.

Bossewitch, J. and Sinnreich, A. (2013) 'The end of forgetting: strategic agency beyond the panopticon', *New Media and Society*, 15(2), pp. 224–242.

boyd, D. and Crawford, K. (2012) 'Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon', *Information, Communication & Society*, 15(5), pp. 662–679.

Brasel, S. A. and Gips, J. (2011) 'Media multitasking behavior: concurrent television and computer usage', *Cyberpsychology, Behavior, and Social Networking*, 14(9), pp. 527–534.

Bruell, A. (2017) 'Ad giants are under pressure to streamline complex structures', *The Wall Street Journal* (English edition), 28 August [online]. Available at: https://www.wsj.com/articles/ad-giants-are-under-pressure-to-streamline-complex-structures-1503943756 (Accessed: 26 January 2021)

Bryman, A. (2016) *Social research methods.* Oxford: Oxford University Press.

Bucher, T. (2017) 'The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms', *Information, Communication & Society*, 20(1), pp. 30–44.

Calder, B. J., Malthouse, E. C. and Schaedel, U. (2009) 'An experimental study of the relationship between online engagement and advertising effectiveness', *Journal of Interactive Marketing*, 23(4), pp. 321–331.

Campbell, J. E. and Carlson, M. (2002) 'Panopticon.com: online surveillance and the commodification of privacy', *Journal of Broadcasting & Electronic Media*, 46(4), pp. 586–606.

Carah, N. and Brodmerkel, S. (2020) 'Critical perspectives on brand culture in the era of participatory and algorithmic media', *Sociology Compass*, 14(2), doi:10.1111/soc4.12752.

Caraway, B. (2011) 'Audience labor in the new media environment: a Marxian revisiting of the audience commodity', *Media, Culture & Society*, 33(5), pp. 693–708.

Carmi, E. (2020) *Media distortions: understanding the power behind spam, noise, and other deviant media. E-book library* [online]. Available at: https://library.oapen.org/handle/20.500.12657/42373 (Accessed: 10 February 2021)

Carmi, E., Yates, S. J., Lockley, E. and Pawluczuk, A. (2020) 'Data citizenship: rethinking data literacy in the age of disinformation, misinformation, and malinformation', *Internet Policy Review*, 9(2), pp. 1–22.

Castells, M. (2007) 'Communication, power and counter-power in the network society', *International Journal of Communication*, 1(1), pp.238–266.

Castells, M. (2013) *Communication power*. New York: Oxford University Press.

Cheney-Lippold, J. (2011) 'A new algorithmic identity: soft biopolitics and the modulation of control', *Theory, Culture & Society*, 28(6), pp. 164–181.

Cialdini, R. B. (2007) *Influence: the psychology of persuasion*, revised edition. New York: HarperCollins.

Ciochetto, L. (2013) *Globalisation and advertising in emerging economies: Brazil, Russia, India and China*. New York: Routledge.

Clapp, R. (2018) *TV adspend rising strongly in India but total share is down* [Online]. Available at: https://www.warc.com/content/article/warc-datapoints/tv_adspend_rising_strongly_in_india_but_total_share_is_down/124230 (Accessed: 23 April 2019)

Clarke, R. (1988) 'Information technology and dataveillance', *Communications of the ACM*, 31(5), pp. 498–512.

Crouch, M. and McKenzie, H. (2006) 'The logic of small samples in interview-based qualitative research', *Social Science Information*, 45(4), pp.483–499.

Dandeker, C. (1990) *Surveillance, power and modernity: bureaucracy and discipline from 1700 to the present day.* Oxford: Polity.

Daniel, C. and Palmer, M. (2007) 'Google's goal: to organise your daily life', *Financial Times,* 22 May [online]. Available at: https://www.ft.com/content/c3e49548-088e-11dc-b11e-000b5df10621 (Accessed: 22 December 2020)

Davies, H. C. (2018) 'Learning to Google: understanding classed and gendered practices when young people use the internet for research', *New Media & Society,* 20(8), pp. 2764–2780.

Davies, S.G. (1997) 'Re-engineering the right to privacy: how privacy has been transformed from a right to a commodity', in Agre, P.E. and Rotenberg, M. (eds.) *Technology and privacy: the new landscape.* Cambridge, MA: MIT Press, pp. 143–165.

Davisson, A. and Booth, P. (eds.) (2017) *Controversies in digital ethics*. New York: Bloomsbury Publishing USA.

DeVito, M. A., Birnholtz, J., Hancock, J.T., French, M. and Liu, S. (2018) How people form folk theories of social media feeds and what it means for how we study self-presentation. *Proceedings of the 2018 CHI conference on human factors in computing systems*, *April 2018,* New York: Association for Computing Machinery, pp. 1-12.

Dixit, R. (2016) *Think big, act small: mobile video advertising in India and Indonesia, Google* [Online]. Available at: https://www.thinkwithgoogle.com/intl/en-apac/marketing-strategies/video/think-big-act-small-mobile-video-advertising-in-india-and-indonesia/ (Accessed: 26 April 2017)

Domo. (2020) *Media usage in an internet minute as of August 2020* [Online]. Available at: https://www.statista.com/statistics/195140/new-user-generated-content-uploaded-by-users-per-minute (Accessed: 01 February 2021)

Edmonds, W.A. and Kennedy, T.D. (2012) *An applied reference guide to research designs: quantitative, qualitative, and mixed methods.* London: SAGE Publications.

Einstein, M. (2017) *Advertising: what everyone needs to know®*. New York: Oxford University Press.

Ellis-Petersen, H. (2020) 'India bans TikTok after Himalayan border clash with Chinese troops', *The Guardian*, 29 June [online]. Available at: https://www.theguardian.com/world/2020/jun/29/india-bans-tiktok-after-himalayan-border-clash-with-chinese-troops (Accessed: 26 November 2020)

Elmer, G. (2003) *Profiling machines: mapping the personal information economy*. Cambridge, MA: MIT Press.

Ernst & Young (2019) *A billion screens of opportunity: India's media and entertainment sector* [Online]. Available at: https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/tmt/ey-how-a-billion-screens-can-turn-india-into-a-m-and-e-powerhouse.pdf (Accessed: 24 April 2020)

Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K. and Kirlik, A. (2016) First I "like" it, then I hide it: folk theories of social feeds. *Proceedings of the 2016 CHI conference on human factors in computing systems, May 2016,* New York: Association for Computing Machinery, pp. 2371–2382.

Facebook (2019) *Find your audience* [Online]. Available at: https://www.facebook.com/business/ads/ad-targeting (Accessed: 31 January 2019)

Findlay, S. (2020) 'India bans TikTok and dozens of other Chinese mobile apps over privacy fears', *Financial Times*, 30 June [online]. Available at: https://www.ft.com/content/08e15c26-48e0-4540-a040-1a8782e84f2e (Accessed: 26 November 2020)

Fiore, Q. and McLuhan, M. (1967) *The medium is the massage*. New York: Penguin Random House.

Flyverbom, M. (2019) *The digital prism: transparency and managed visibilities in a datafied world*. Cambridge: Cambridge University Press.

Fotopoulou, A. (2018) 'From networked to quantified self: self-tracking and the moral economy of data sharing', in Papacharissi, Z. (ed.) *A networked self and platforms, stories, connections.* New York: Routledge, pp. 144–159.

Foucault, M. (2008) '"Panopticism" from "discipline & punish: the birth of the prison"', *Race/Ethnicity: Multidisciplinary Global Contexts*, 2(1), pp. 1–12. http://www.jstor.org/stable/25594995

Fuchs, C. (2012) 'Dallas Smythe today – the audience commodity, the digital labour debate, Marxist political economy and critical theory', *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 10(2), pp. 692–740.

Ganascia, J. G. (2010) 'The generalized sousveillance society'. *Social Science Information*, 49(3), pp. 489–507.

Gandy Jr, O. H. (1989) 'The surveillance society: information technology and bureaucratic social control', *Journal of Communication*, 39(3), pp. 61–76.

Gandy Jr, O. H. (1993) *The panoptic sort: a political economy of personal information*. New York: Oxford University Press.

Gandy, R. (1996) 'Human versus mechanical intelligence'. *Machines and Thought: The Legacy of Alan Turing*, Oxford: Clarendon, pp. 125–136.

Georgakopoulou, A. and Spilioti, T. (eds). (2015) *The Routledge handbook of language and digital communication*. Abingdon: Routledge.

Gillespie, T. (2014) 'The relevance of algorithms' in Gillespie, T., Boczkowski, P. J. and Foot, K. A. (eds.) *Media technologies: essays on communication, materiality, and society*. Cambridge, MA: MIT Press, p. 167.

Gilliom, J. (2001) *Overseers of the poor: surveillance, resistance, and the limits of privacy*. Chicago: University of Chicago Press.

Glass, R. and Callahan, S. (2014) *The big data-driven business: how to use big data to win customers, beat competitors, and boost profits*. Hoboken, New Jersey: John Wiley & Sons.

Golafshani, N. (2003) 'Understanding reliability and validity in qualitative research', *The Qualitative Report*, 8(4), pp. 597–607.

Google, 2017 consumer barometer [Online]. Available at: https://www.consumerbarometer.com/en/graphbuilder/?question=M1&filter=country:india,united_kingdom,united_states (Accessed: 24 April 2018)

Goyal, M. (2020) '25 years of internet in India and how it has transformed our lives',' *The Economic Times,* 16 August [online]. Available at: https://economictimes.indiatimes.com/tech/internet/the-india-story-after-25-years-of-internet-and-how-it-has-transformed-our-lives/articleshow/77565596.cms (Accessed: 22 October 2020)

Green, A. (2017) *Audience measurement 5.0 – pushing the boundaries* [Online]. Available at: https://www.ipsos.com/en/audience-measurement-50-pushing-boundaries (Accessed: 23 April 2019)

Grimshaw, C. (2017) *A post-demographic age*. [Online]. Available at: https://www.warc.com/content/article/a-post-demographic-age/en-gb/117252 (Accessed: 27 May 2018)

Gurses, S., van Hoboken, J. V. J. (2017) *Privacy after the agile turn* [Online]. Available at: https://osf.io/preprints/socarxiv/9gy73/ (Accessed: 23 April 2021)

Haggin, P. and Vranica, S. (2021) 'Apple's moves vex digital advertisers', *The Wall Street Journal* (English edition), 09 June [online]. Available at: https://www.wsj.com/articles/apples-moves-to-tighten-flow-of-user-data-leave-advertisers-anxious-11623189903 (Accessed: 26 June 2021)

Hartman, T., Kennedy, H., Steedman, R. and Jones, R. (2020) 'Public perceptions of good data management: findings from a UK-based survey', *Big Data and Society*, 7(1). doi: 10.1177/2053951720935616

Hepp, A., Breiter, A. and Friemel, T. N. (2018) 'Digital traces in context | digital traces in context—an introduction', *International Journal of Communication*, 12, p.11.

Humphreys, L. (2011) 'Who's watching whom? a study of interactive technology and surveillance', *Journal of Communication,* 61(4), pp. 575–595.

Iliadis, A. and Russo, F. (2016) 'Critical data studies: an introduction', *Big Data & Society*, 3(2), p. 2053951716674238.

Indian Parliament (2019) *Personal data protection bill 2019* [Online]. Available at: http://164.100.47.4/BillsTexts/LSBillTexts/Asintroduced/373_2019_LS_Eng.pdf(Accessed: 23 April 2021)

Jain, S. (2015) 'India: multiple media explosions', in Nordenstreng, K. and Thussu, D.K. (eds.) *Mapping BRICS media.* New York: Routledge, pp. 145–165.

James, C. (2009) *Young people, ethics, and the new digital media: a synthesis from the GoodPlay project*. Cambridge, MA: MIT Press, p.128.

Jenkins, H. (2006) *Convergence culture*. New York: New York University Press.

Johnson, R. B. (1997) 'Examining the validity structure of qualitative research', *Education,* 118(2), pp. 282–293.

Jowett, G. S. and O'donnell, V. (2018) *Propaganda & Persuasion*. California: SAGE Publications.

Kandias, M., Mitrou, L., Stavrou, V. and Gritzalis, D. (2013) *Which side are you on? a new panopticon vs. privacy* [Online]. Available at: https://ieeexplore.ieee.org/document/7223159 (Accessed: 23 January 2019)

Kantchev, G. and Gross, J. (2018) 'Cambridge university pushes back against Facebook scrutiny of research apps', *The Wall Street Journal* (English edition), 14 April [online]. Available at: https://www.wsj.com/articles/cambridge-university-pushes-back-against-facebook-scrutiny-of-research-apps-1523703601 (Accessed: 26 March 2020)

Kaplan, M. (2020) ‘The self-consuming commodity: audiences, users, and the riddle of digital labor’, *Television & New Media*, 21(3), pp. 240–259.

Kennedy, H. (2016) *Post, mine, repeat: social media data mining becomes ordinary*. London: Springer.

Kennedy, H. (2018) ‘Living with data: aligning data studies and data activism through a focus on everyday experiences of datafication’, *Krisis: Journal for Contemporary Philosophy,* 2018(1), pp. 18–30.

Kitchin, R. (2014) *The data revolution: big data, open data, data infrastructures and their consequences.* London: SAGE Publications.

Kitchin, R. and Dodge, M. (2014) *Code/space: software and everyday life*. Cambridge, MA: MIT Press.

Klosowski, T. (2021) ‘The state of consumer data privacy laws in the US (and why it matters)’, *The New York Times Wirecutter,* 06 September [online]. Available at: https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us (Accessed: 09 September 2021)

Kohli-Khandekar, V. (2013) *The Indian media business*. New Delhi: SAGE Publications India.

Lee, M. (2011) ‘Google ads and the blindspot debate’, *Media, Culture & Society*, 33(3), pp. 433–447.

Levy, S. (2021) *In the plex: how Google thinks, works, and shapes our lives.* New York: Simon & Schuster.

Lomborg, S. and Kapsch, P. H. (2020) ‘Decoding algorithms’, *Media, Culture & Society,* 42(5), pp. 745–761.

Lupton, D. (2016) *The quantified self*. Cambridge: Polity.

Lupton, D. (2018) 'How do data come to matter? Living and becoming with personal data'. *Big Data & Society*, 5(2), doi: 10.1177/2053951718786314.

Lupton, D. (2020a) ‘Better understanding about what's going on’: young Australians’ use of digital technologies for health and fitness’, *Sport, Education and Society*, 25(1), pp. 1–13.

Lupton, D. (2020b) *Data selves: more-than-human perspectives*. London: Polity.

Lupton, D. (2021) ‘‘Not the real me’: social imaginaries of personal data profiling’, *Cultural Sociology,* 15(1), pp. 3–21.

Lyon, D. (2014) ‘Surveillance, Snowden, and big data: capacities, consequences, critique’, *Big Data & Society*, 1(2), p.2053951714541861.

Masih, N. (2020) *India's TikTok ban eats into his empire* [Online]. Available at: https://www.pressreader.com/usa/the-washington-post/20200711/281956020080738 (Accessed: 23 September 2020)

Mathiesen, T. (1997) 'The viewer society: Michel Foucault's panopticon' revisited'. *Theoretical criminology*, 1(2), pp. 215–234.

Mayernik, M. S. and Acker, A. (2018) ‘Tracing the traces: the critical role of metadata within networked communications’, *Journal of the Association for Information Science and Technology*, 69(1), pp. 177–180.

Mayer-Schönberger, V. (2009) *Delete: the virtue of forgetting in the digital age. E-book library* [online]. Available at: https://www.proquest.com/docview/1681853783 (Accessed: 20 October 2020)

Mayer-Schönberger, V. and Cukier, K. (2013) *Big data: a revolution that will transform how we live, work, and think*. New York: Houghton Mifflin Harcourt.

McDonald, J. and Clapp, R. (2020) *Global ad trends: state of the industry 2020/21* [Online]. Available at: https://www.warc.com/content/article/warc-data/global-ad-trends-outlook-2020-21/134804 (Accessed: 23 June 2020)

McKinnon, J. D. (2020) ‘TikTok shutdown argued before federal judge’, *The Wall Street Journal* (English edition)*,* 04 November [online]. Available at: https://www.wsj.com/articles/tiktok-shutdown-argued-before-federal-judge-11604518470 (Accessed: 20 January 2021)

McStay, A. (2009) *Digital advertising*. New York: Palgrave Macmillan.

McStay, A. (2017) 3.1 Micro-Moments, liquidity, intimacy and automation: developments in programmatic ad-tech in Siegert, G., Rimscha, M. and Grubenmann, S. (eds.) *Commercial communication in the digital age: information or disinformation?* Berlin, Boston: De Gruyter Saur, pp. 143–160. doi: 10.1515/9783110416794-009

McStay, A. (2018) 'Digital advertising and adtech: programmatic platforms, identity and moments', in Powell, H., MacRury, I. and Hardy, J. (eds.) *The advertising handbook.* New York: Routledge, pp. 88–100.

Meehan, E. R. (1984) 'Ratings and the institutional approach: a third answer to the commodity question', *Critical Studies in Mass Communication*, 1(2), pp. 216–25.

Meehan, E. R. and Torre, P. J. (2011) 'Markets in theory and markets in television' in Wasko, J., Murdock, G. and Sousa, H. (eds). *The handbook of political economy of communications*. Oxford: Wiley Blackwell, pp. 62-82. doi: 10.1002/9781444395402.ch3

Meeker, M. (2019) *Internet trends 2019* [Online]. Available at: https://www.bondcap.com/report/itr19/ (Accessed: 01 February 2021)

Meyrowitz, J. (2009) 'We liked to watch: television as progenitor of the surveillance society'. *The Annals of the American Academy of Political and Social Science*, 625(1), pp. 32–48.

Napoli, P. M. (2014) 'Automated media: an institutional theory perspective on algorithmic media production and consumption', *Communication Theory*, 24(3), pp. 340–360.

Nguyen, P. and Solomon, L. (2018) *Consumer data and the digital economy* [Online]. Available at: https://cprc.org.au/app/uploads/2018/07/Consumer-Data-and-the-Digital-Economy_smallest-file-size.pdf (Accessed: 20 February 2021)

Nielsen, R. K. (2016) 'Folk theories of journalism: the many faces of a local newspaper', *Journalism Studies*, 17(7), pp. 840–848.

Ninan, T. N. (2018) 'The political economy of reforms: the art of the possible' in Mohan, R. (ed.) *India transformed: twenty-five years of economic reforms.* Washington, D.C.: Brookings Institution Press, pp. 77–90.

Nord VPN (2020) *National privacy test* [Online]. Available at: https://s1.nordcdn.com/nord/misc/0.34.0/vpn/brand/national-privacy-test-report.pdf (Accessed: 06 September 2021)

Nordenstreng, K. and Thussu, D. K. (eds.) (2015) *Mapping BRICS media*. New York: Routledge.

O. Nyumba, T., Wilson, K., Derrick, C. J. and Mukherjee, N. (2018) 'The use of focus group discussion methodology: insights from two decades of application in conservation', *Methods in Ecology and Evolution*, 9(1), pp. 20–32.

O'Neil, C. (2016*) Weapons of math destruction*. London: Penguin Books.

Palmer, R. A. (2019) 'A "deep story" about American journalism: using "episodes" to explore folk theories of journalism', *Journalism Studies*, 20(3), pp. 327–344.

Pangrazio, L. and Selwyn, N. (2019) ''Personal data literacies': a critical literacies approach to enhancing understandings of personal digital data', *New Media & Society*, 21(2), pp. 419–437.

Papacharissi, Z. (ed.) (2010) *A networked self: identity, community, and culture on social network sites*. New York: Routledge.

Park, M. J. and Curran, J. (eds.) (2000) *De-Westernizing media studies*. New York: Routledge.

Pasquale, F. (2015) *The black box society: the secret algorithms that control money and information.* Cambridge, MA and London, England: Harvard University Press. https://doi.org/10.4159/harvard.9780674736061

Paul, S. L. and Wang, G. (1995) 'Satellite TV in Asia: forming a new ecology', *Telecommunications Policy*, 19(2), pp. 135–149.

Phillips, D. J. (2004) 'Privacy policy and PETs: The influence of policy regimes on the development and social implications of privacy enhancing technologies', *New Media & Society*, 6(6), pp. 691–706. https://doi-org.soton.idm.oclc.org/10.1177/146144804042523.

Pine, B. J. and Gilmore, J. H. (1998) 'Welcome to the experience economy', *Harvard Business Review*, 76, pp. 97–105.

Pink, S., Postill, J., Mackley, K.L. and Astari, N. (2017) 'Digital-visual stakeholder ethnography', *Sociological Research Online*, (4), pp. 174–192. doi: 10.1177/1360780417726736.

Protess, D. and McCombs, M. E. (2016) *Agenda setting: readings on media, public opinion, and policymaking.* New York: Routledge.

Pybus, J., Coté, M. and Blanke, T. (2015) 'Hacking the social life of big data', *Big Data & Society*, 2(2), p.2053951715616649.

Rabiee, F. (2004) 'Focus-group interview and data analysis', *Proceedings of The Nutrition Society*, 63(4), pp. 655–660.

Raley, R. (2013) 'Dataveillance and countervailance' in Gitelman, L. (ed.) *'Raw data' is an oxymoron*. Cambridge, MA: MIT Press, pp. 121–146.

Reynolds, T. J. (2006) 'Methodological and strategy development implications of decision segmentation', *Journal of Advertising Research*, 46(4), pp. 445–461.

Rip, A. (2006) 'Folk theories of nanotechnologists', *Science as Culture*, 15(4), pp. 349–365.

Rogers, R. (2009) *Post-demographic machines: studying social networking* [Online]. Available at: http://www.govcom.org/publications/full_list/WalledGarden_ch04_RR.pdf (Accessed: 25 November 2018)

Rossiter, J. R. and Danaher, P. J. (1998) *Advanced media planning Vol. (1).* New York: Springer.

Roy, R. and Li, S. (2020) 'India bans TikTok, dozens of other Chinese apps after border clash; UC Browser and WeChat are among the banned mobile apps*', The Wall Street Journal* (English edition)*,* 30 June [online]. Available at: https://www.wsj.com/articles/india-blocks-dozens-of-chinese-apps-including-tiktok-following-border-clash-11593447321 (Accessed: 15 July 2020)

Ruckenstein, M. and Granroth, J. (2020) 'Algorithms, advertising and the intimacy of surveillance', *Journal of Cultural Economy,* 13(1), pp. 12–24.

S.Wall, D. (2019) 'Surveillant internet technologies and the growth in information capitalism: spams and public trust in the information society' in Haggerty, K. and Ericson, R. (eds.) *The new politics of surveillance and visibility*. Toronto: University of Toronto Press, pp. 340–362. doi: 10.3138/9781442681880-015

Sadowski, J. (2020) *Too smart: how digital capitalism is extracting data, controlling our lives, and taking over the world.* Cambridge, MA: MIT Press.

Sandel, M. J. (2012) *What money can't buy: The moral limits of markets.* New York: Farrar, Straus and Giroux.

Saulles, M. D. and Horner, D. S. (2011) 'The portable panopticon: morality and mobile technologies', *Journal of Information, Communication & Ethics in Society*, 9(3), pp. 206–216.

Schechner, S., Grind, K. and Haggin, P. (2020) 'Tech firms are spying on you. In a pandemic, governments say that's OK', *The Wall Street Journal* (English edition)*,* 15 June [online]. Available at: https://www.wsj.com/articles/once-pariahs-location-tracking-firms-pitch-themselves-as-covid-sleuths-11592236894 (Accessed: 11 July 2020)

Shklovski, I., Mainwaring, S.D., Skúladóttir, H.H. and Borgthorsson, H. Leakiness and creepiness in app space: perceptions of privacy and mobile app use. *Proceedings of the SIGCHI conference on human factors in computing systems, April 2018,* Montreal: pp. 2347–2356.

Shoenberger, H. (2017) 'Targeted digital advertising and privacy' in Rodgers, S. and Thorson, E. (eds.) *Digital advertising: theory and research.* New York: Routledge, pp. 300–309.

Sissors, J. Z. and Goodrich W. B. (1996) *Media planning workbook*. Lincolnwood: NTC/Contemporary Publishing Group.

Skeggs, B. and Yuill, S. (2016a) 'Capital experimentation with person/a formation: how Facebook's monetization refigures the relationship between property, personhood and protest', *Information, Communication & Society,* 19(3), pp. 380–396.

Skeggs, B. and Yuill, S. (2016b) 'The methodology of a multi-model project examining how Facebook infrastructures social relations', *Information, Communication & Society*, 19(10), pp. 1356–1372.

Smythe, D. W. (1981) 'On the audience commodity and its work' in Durham, M. G and Kellner, D. M. (eds.) *Media and cultural studies: keyworks*. Oxford: Wiley-Blackwell, pp. 230–256.

Solove, D. (2004) *The digital person: technology and privacy in the information age*. New York: New York University Press.

Spindle, B. (2019) ''The youth bulge': India struggles to employ an exploding population', *The Wall Street Journal* (English edition), 22 April [online]. Available at: https://www.wsj.com/articles/the-youth-bulge-india-struggles-to-employ-an-exploding-population-11555925401 (Accessed: 26 August 2020)

Srinivasan, S., Rutz, O. J. and Pauwels, K. (2016) 'Paths to and off purchase: quantifying the impact of traditional marketing and online consumer activity', *Journal of the Academy of Marketing Science*, 44(4), pp. 440–453.

Steedman, R., Kennedy, H. and Jones, R. (2020) 'Complex ecologies of trust in data practices and data-driven systems', *Information, Communication & Society,* 23(6), pp. 817–832.

Stokes, J. (2012) *How to do media and cultural studies*. London: SAGE Publications.

Straubhaar, J., LaRose, R. and Davenport, L. (2013) *Media now: understanding media, culture, and technology*. Boston, MA: Wadsworth Cengage Learning.

Swisher, K. (2019) 'We're not going to take it anymore', *The New York Times,* 10 April [online]. Available at: https://www.nytimes.com/2019/04/10/opinion/internet-privacy-regulation.html (Accessed: 26 March 2020)

Szulc, L. (2018) 'Profiles, identities, data: making abundant and anchored selves in a platform society', *Communication Theory*, 29(3), pp. 169–188

Tau, B. (2020) 'Next step in government data tracking is the internet of things', *The Wall Street Journal* (English edition), 27 November [online]. Available at: https://www.wsj.com/articles/next-step-in-government-data-tracking-is-the-internet-of-things-11606478401 (Accessed: 02 January 2021)

The Great Hack (2019) Directed by Jehane Noujaim and Karim Ame [Film]. California: Netflix.

The Nielsen Company (2018) *The Nielsen CMO Report 2018* [Online]. Available at: https://www.nielsen.com/us/en/insights/reports/2018/cmo-report-2018-digital-media-roi-measurement-omnichannel-marketing-technology.html (Accessed: 10 March 2019)

The Wall Street Journal (2021) *Inside TikTok's highly secretive algorithm*. 21 July [online]. Available at: https://www.wsj.com/video/series/inside-tiktoks-highly-secretive-algorithm/investigation-how-tiktok-algorithm-figures-out-your-deepest-desires/6C0C2040-FF25-4827-8528-2BD6612E3796 (Accessed: 27 July 2021)

Thompson, D. (2010) *Google's CEO: 'The laws are written by lobbyists'* Available at: https://www.theatlantic.com/technology/archive/2010/10/googles-ceo-the-laws-are-written-by-lobbyists/63908 (Accessed: 22 December 2020)

Thrift, N. (2005) *Knowing capitalism*. London: SAGE Publications.

Toff, B. and Nielsen, R. K. (2018) '"I just Google it": folk theories of distributed discovery', *Journal of Communication*, 68(3), pp. 636–657.

Tsesis, A. (2018) 'Marketplace of ideas, privacy, and the digital audience', *Notre Dame Law Review*, 94, p.1585.

Turow, J. (2005) 'Audience construction and culture production: marketing surveillance in the digital age', *The ANNALS of the American Academy of Political and Social Science*, 597(1), pp. 103–121. doi: 10.1177/0002716204270469.

Turow, J. (2011) *The daily you: how the new advertising industry is defining your identity and your worth*. New Haven: Yale University Press.

Turow, J., McGuigan, L. and Maris, E.R. (2015) 'Making data mining a natural part of life: physical retailing, customer surveillance and the 21st century social imaginary', *European Journal of Cultural Studies*, 18(4-5), pp. 464–478.

Van Dijck, J. (2013) *The culture of connectivity: a critical history of social media*. New York: Oxford University Press.

Van Dijck, J. and Poell, T. (2013) 'Understanding social media logic', *Media and Communication,* 1(1), pp. 2–14.

Varis, P. (2016) 'Digital ethnography' in Georgakopoulou, A. and Spilioti, T. (eds). *The Routledge handbook of language and digital communication.* Abingdon: Routledge, pp. 55–68.

Vega, T. (2012) 'Voters say they are wary of ads made just for them', *The New York Times,* 23 July [online]. Available at: https://www.nytimes.com/2012/07/24/business/media/survey-shows-voters-are-wary-of-tailored-political-ads.html (Accessed: 26 March 2020)

Vranica, S. (2018a) 'Facebook and Google face emboldened antagonists: big advertisers', *The Wall Street Journal* (English edition), 25 March [online]. Available at: https://www.wsj.com/articles/facebook-and-google-face-emboldened-antagonists-big-advertisers-1521998394 (Accessed: 02 January 2019)

Vranica, S. (2018b) 'Unilever threatens to reduce ad spending on tech platforms that don't combat divisive content', *The Wall Street Journal* (English edition)*,* 11 February [online]. Available at: https://www.wsj.com/articles/unilever-threatens-to-reduce-ad-spending-on-tech-platforms-that-dont-combat-divisive-content-1518398578 (Accessed: 12 January 2019)

Wall, D.S. (2019) 'Surveillant internet technologies and the growth in information capitalism: spams and public trust in the information society', in Haggerty, K. and Ericson, R. (eds.) *The new politics of surveillance and visibility*. Toronto: University of Toronto Press, pp. 340–362. https://doi.org/10.3138/9781442681880-015

WARC (2019) *Adspend database* [Online]. Available at: https://www.warc.com/data/adspend/takeaway-tables (Accessed: 23 April 2019)

Warner, C., Lederer, W. A. and Moroz, B. (2020) *Media selling: digital, television, audio, print and cross-platform.* Oxford: Wiley-Blackwell.

Watkins, E. A. (2019) *Guide to advertising technology* [Online]. Available at: https://academiccommons.columbia.edu/doi/10.7916/d8-9aep-r971 (Accessed: 23 April 2020)

Webster, J., Phalen, P. and Lichty, L. (2013) *Ratings analysis: audience measurement and analytics*. New York: Routledge.

West, E. (2019) 'Amazon: surveillance as a service'. *Surveillance & Society*, 17(1–2), pp. 27–33.

Westin, A. F. (2000) 'Intrusions', *Public Perspective*, 11(6), pp. 8–11.

Whitaker, R. (2000) *The end of privacy: how total surveillance is becoming a reality.* New York: The New Press.

Whitson, J. R. (2013) 'Gaming the quantified self'. *Surveillance & Society*, 11(1–2), pp. 163–176.

Wolf, M. J. (1999) *The entertainment economy. The mega-media forces that are re-shaping our lives*. New York: Penguin Random House.

Wolff, M. (2017) *Television is the new television: the unexpected triumph of old media in the digital age*. New York: Penguin Random House.

Wu, T. (2017) *The attention merchants: the epic scramble to get inside our heads*. New York: Vintage Books.

Young, A. (2014) *Brand media strategy: integrated communications planning in the digital era*. New York: Palgrave Macmillan.

Young, M. (2017) *Ogilvy on advertising in the digital age*. London: Goodman Books.

Ytre-Arne, B. and Moe, H. (2020) 'Folk theories of algorithms: Understanding digital irritation', *Media, Culture & Society*. doi:10.1177/0163443720972314.

Zackon, H., Marci, C. and Nuta, N. (2017) *The mind of a reader* [Online]. Available at: http://www.researchexcellence.com/files/pdf/2017-03/id423_cre_the_mind_of_the_viewer_arf_presentation_3.14.17.pdf (Accessed: 24 November 2018)

Zuboff, S. (2019) *The age of surveillance capitalism: the fight for a human future at the new frontier of power.* New York: Public Affairs.

Bibliography

Achrol, R.S. and Kotler, P. (2012) 'Frontiers of the marketing paradigm in the third millennium', *Journal of the Academy of Marketing Science*, 40(1), pp. 35–52.

Agarwal, S. (*2017) Internet users projected to cross 450 million by June 2017: IAMAI* [Online]. Available at: https://cio.economictimes.indiatimes.com/news/internet/internet-users-to-cross-450-million-by-june-2017-iamai/57413100 (Accessed: 23 April 2019)

Agrawal, R. (2018) *India connected: how the smartphone is transforming the world's largest democracy*. New York: Oxford University Press.

Almquist, E., Senior, J. and Bloch, N. (2016) 'The elements of value. (cover story)', *Harvard Business Review*, 94(9), pp. 46–92.

Ambler, T. (2000) 'Persuasion, pride and prejudice: how ads work', *International Journal of Advertising*, 19(3), pp. 299–315.

Ambler, T. and Burne, T. (1999) 'The impact of affect on memory of advertising', *Journal of Advertising Research*, 39, pp. 25–34.

Anand, B.N. and Khanna, T. (2000) *Must Zee TV* [Online]. Available at: https://www.hbs.edu/faculty/Pages/item.aspx?num=27229 (Accessed: 27 May 2019)

Ananny, M. and Crawford, K. (2018) 'Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability', *New Media & Society*, 20(3), pp. 973–989.

Anderson, C. (2009) 'Essay: the economics of giving it away', *The Wall Street Journal* (English edition), 31 January [online]. Available at: https://www.wsj.com/articles/SB123335678420235003 (Accessed: 26 August 2020)

Andrejevic, M. (2007) *iSpy: surveillance and power in the interactive era.* Lawrence: University Press of Kansas.

Andrejevic, M. (2013) *Infoglut: how too much information is changing the way we think and know*. New York: Routledge.

Andrejevic, M. (2014) 'Big data, big questions| the big data divide', *International Journal of Communication,* 8, p.17.

Andrejevic, M. (2019) 'Automating surveillance', *Surveillance & Society*, 17(1/2), pp. 7–13.

Bibliography

Andrejevic, M., Hearn, A. and Kennedy, H. (2015) 'Cultural studies of data mining: introduction', *European Journal of Cultural Studies*, 18(4-5) pp. 379–394.

Arora, P. (2016) 'Bottom of the data pyramid: big data and the global south', *International Journal of Communication,* 10, p.19.

Ashton, D. and Gowland-Pryde, R. (2019) 'Arts audience segmentation: data, profiles, segments and biographies', *Cultural Trends*, *28*(2-3), pp. 146–161.

Assael, H. (2011) 'From silos to synergy: A fifty-year review of cross-media research shows synergy has yet to achieve its full potential', *Journal of Advertising Research,* 51 (150th Anniversary Supplement), pp. 42–58.

Athique, A. (2012) *Indian media*. Cambridge: Polity.

Auletta, K. (2018) *Frenemies: the epic disruption of the ad business (and everything else)*. London: Penguin Books.

Bakir, V. and McStay, A. (2018) 'Fake news and the economy of emotions: Problems, causes, solutions', *Digital Journalism*, 6(2), pp. 154–175.

Batra, A. (ed.) (2003) *Marketing whitebook 2003–2004.* New Delhi: BWBooks.

Batra, A. (ed.) (2017) *Marketing whitebook 2017–2018.* New Delhi: BWBooks.

Batra, R. and Ray, M.L. (1985) 'How advertising works at contact', *Psychological processes and advertising effects*, pp. 13–44.

Batra, R., Ahuvia, A. and Bagozzi, R.P. (2012) 'Brand love', *Journal of Marketing*, 76(2), pp. 1–16.

Bawden, D. (2008) 'Origins and concepts of digital literacy' in Lankshear, C. and Knobel, M. (eds.) *Digital literacies: concepts, policies and practices*, 30(2008), New York: Peter Lang, pp. 17–32.

Beer, D. (2009) 'Power through the algorithm? Participatory web cultures and the technological unconscious', N*ew Media & Society*, 11(6), pp. 985–1002.

Bermejo, F. (2009) 'Audience manufacture in historical perspective: from broadcasting to Google', *New Media & Society*, 11(1-2), pp. 133–154.

Berry, D. (2011) *The philosophy of software: code and mediation in the digital age*. *E-book library* [online]. Available at: http://lib.myilibrary.com/?id=306769 (Accessed 14 March 2021)

Binet, L. and Field, P. (2009) 'Empirical generalizations about advertising campaign success', *Journal of Advertising Research*, 49(2), pp. 130–133.

Bodle, R. (2016) 'A critical theory of advertising as surveillance' in Hamilton, J. F., Bodle, R. and Korin, E. (eds.) *Explorations in critical studies of advertising*. New York: Routledge, pp. 138–152.

Bogard, W. (1996) *The simulation of surveillance: hyper-control in telematic societies*. Cambridge: Cambridge University Press.

Bogart, L. (1994) 'Who pays for the media?', *Journal of Advertising Research*, 34(2), pp. 11–19.

Bolin, G. and Andersson Schwarz, J. (2015) 'Heuristics of the algorithm: big data, user interpretation and institutional translation', *Big Data & Society*, 2(2), pp. 1–12.

Borges, J.L. (1964) *Labyrinths: selected stories & other writings.* New York: New Directions Publishing.

Bossewitch, J. and Sinnreich, A. (2013) 'The end of forgetting: strategic agency beyond the panopticon', *New Media and Society*, 15(2), pp. 224–242.

boyd, D. and Crawford, K. (2012) 'Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon', *Information, Communication & Society*, 15(5), pp. 662–679.

Braidotti, R. (2019) 'A theoretical framework for the critical posthumanities', *Theory, Culture & Society*, 36(6), pp. 31–61.

Brasel, S. A. and Gips, J. (2011) 'Media multitasking behavior: concurrent television and computer usage', *Cyberpsychology, Behavior, and Social Networking*, 14(9), pp. 527–534.

Bruell, A. (2017) 'Ad giants are under pressure to streamline complex structures', *The Wall Street Journal* (English edition), 28 August [online]. Available at: https://www.wsj.com/articles/ad-giants-are-under-pressure-to-streamline-complex-structures-1503943756 (Accessed: 26 January 2021)

Bryman, A. (2016) *Social research methods.* Oxford: Oxford University Press.

Bucher, T. (2017) 'The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms', *Information, Communication & Society*, 20(1), pp. 30–44.

Bulger, M. and Davison, P. (2018) 'The promises, challenges, and futures of media literacy', *Journal of Media Literacy Education*, 10(1), pp. 1–21.

Bibliography

Byron, E. (2011) 'In-store sales begin at home - marketers rethink strategies as shoppers research online before compiling lists', *The Wall Street Journal* (English edition), 25 April [online]. Available at: https://www.wsj.com/articles/SB10001424052748703907004576279253672899290 (Accessed: 26 June 2019)

Calder, B. J., Malthouse, E. C. and Schaedel, U. (2009) 'An experimental study of the relationship between online engagement and advertising effectiveness', *Journal of Interactive Marketing*, 23(4), pp. 321–331.

Campbell, J. E. and Carlson, M. (2002) 'Panopticon.com: online surveillance and the commodification of privacy', *Journal of Broadcasting & Electronic Media*, 46(4), pp. 586–606.

Cannon, H. M., Leckenby, J. D. and Abernethy, A. (2002) 'Beyond effective frequency: Evaluating media schedules using frequency value planning', *Journal of Advertising Research*, 42(6), pp. 33–47.

Carah, N. and Brodmerkel, S. (2020) 'Critical perspectives on brand culture in the era of participatory and algorithmic media', *Sociology Compass*, 14(2), doi:10.1111/soc4.12752.

Caraway, B. (2011) 'Audience labor in the new media environment: a Marxian revisiting of the audience commodity', *Media, Culture & Society*, 33(5), pp. 693–708.

Carmi, E. (2020) *Media distortions: understanding the power behind spam, noise, and other deviant media. E-book library* [online]. Available at: https://library.oapen.org/handle/20.500.12657/42373 (Accessed: 10 February 2021)

Carmi, E., Yates, S. J., Lockley, E. and Pawluczuk, A. (2020) 'Data citizenship: rethinking data literacy in the age of disinformation, misinformation, and malinformation', *Internet Policy Review*, 9(2), pp. 1–22.

Carrier, L.M., Cheever, N.A., Rosen, L.D., Benitez, S. and Chang, J. (2009) 'Multitasking across generations: multitasking choices and difficulty ratings in three generations of Americans', *Computers in Human Behavior*, 25(2), pp. 483–489.

Castells, M. (2007) 'Communication, power and counter-power in the network society', *International Journal of Communication*, 1(1), pp.238–266.

Castells, M. (2013) *Communication power*. New York: Oxford University Press.

Chaiken, S. (1987) 'The heuristic model of persuasion' in Zanna, M.P., Olson, J.M. and Herman C.P. (eds.) *Social influence: the Ontario symposium Volume 5*. New York: Psychology Press, pp. 3–39.

Chang, H. (2016) *Autoethnography as method* (Vol. 1). New York: Routledge.

Chang, Y. and Thorson, E. (2004) 'Television and web advertising synergies', *Journal of Advertising*, 33(2), pp. 75–84.

Cheney-Lippold, J. (2011) 'A new algorithmic identity: soft biopolitics and the modulation of control', *Theory, Culture & Society*, 28(6), pp. 164–181.

Cheng, V.T. and Loi, M.K. (2014) 'Handling negative online customer reviews: the effects of elaboration likelihood model and distributive justice', *Journal of Travel & Tourism Marketing*, 31(1), pp. 1–15.

Cheong, Y., De Gregorio, F. and KiM, K. (2010) 'The power of reach and frequency in the age of digital advertising: offine and online media demand different metrics', *Journal of Advertising Research*, 50(4), pp. 403–415.

Cho, C.H. (1999) 'How advertising works on the WWW: modified elaboration likelihood model', *Journal of Current Issues & Research in Advertising*, 21(1), pp. 34–50.

Christensen, C. and Prax, P. (2012) 'Assemblage, adaptation and apps: Smartphones and mobile gaming', *Continuum*, 26(5), pp. 731–739.

Cialdini, R. B. (2007) *Influence: the psychology of persuasion*, revised edition. New York: HarperCollins.

Ciochetto, L. (2013) *Globalisation and advertising in emerging economies: Brazil, Russia, India and China*. New York: Routledge.

Clapp, R. (2018) *TV adspend rising strongly in India but total share is down* [Online]. Available at: https://www.warc.com/content/article/warc-datapoints/tv_adspend_rising_strongly_in_india_but_total_share_is_down/124230 (Accessed: 23 April 2019)

Clarke, R. (1988) 'Information technology and dataveillance', *Communications of the ACM*, 31(5), pp. 498–512.

Cohen, N.S. (2008) 'The valorization of surveillance: towards a political economy of Facebook', *Democratic Communiqué*, 22(1), pp. 5–5.

Bibliography

Collard, A.S., De Smedt, T., Dufrasne, M., Fastrez, P., Ligurgo, V., Patriarche, G. and Philippette, T., (2017) 'Digital media literacy in the workplace: a model combining compliance and inventivity', *Italian Journal of Sociology of Education*, 9(1).

Crouch, M. and McKenzie, H. (2006) 'The logic of small samples in interview-based qualitative research', *Social Science Information*, 45(4), pp.483–499.

Dandeker, C. (1990) *Surveillance, power and modernity: bureaucracy and discipline from 1700 to the present day*. Oxford: Polity.

Daniel, C. and Palmer, M. (2007) 'Google's goal: to organise your daily life', *Financial Times,* 22 May [online]. Available at: https://www.ft.com/content/c3e49548-088e-11dc-b11e-000b5df10621 (Accessed: 22 December 2020)

Davies, H. C. (2018) 'Learning to Google: understanding classed and gendered practices when young people use the internet for research', *New Media & Society,* 20(8), pp. 2764–2780.

Davies, S.G. (1997) 'Re-engineering the right to privacy: how privacy has been transformed from a right to a commodity', in Agre, P.E. and Rotenberg, M. (eds.) *Technology and privacy: the new landscape*. Cambridge, MA: MIT Press, pp. 143–165.

Davisson, A. and Booth, P. (eds.) (2017) *Controversies in digital ethics*. New York: Bloomsbury Publishing USA.

De Montjoye, Y.A., Hidalgo, C.A., Verleysen, M. and Blondel, V.D. (2013) 'Unique in the crowd: The privacy bounds of human mobility', *Scientific Reports*, 3(1), pp. 1–5.

Denzin, N.K. (2006) 'Analytic autoethnography, or déjà vu all over again', *Journal of Contemporary Ethnography*, 35(4), pp. 419–428.

DeVito, M. A., Birnholtz, J., Hancock, J.T., French, M. and Liu, S. (2018) How people form folk theories of social media feeds and what it means for how we study self-presentation. *Proceedings of the 2018 CHI conference on human factors in computing systems*, *April 2018,* New York: Association for Computing Machinery, pp. 1-12.

Dijkstra, M. (2002) *An experimental investigation of synergy effects in multiple-media advertising campaigns*, PhD thesis, Tilburg University, Netherlands, https://pure.uvt.nl/ws/portalfiles/portal/478293/89620.pdf (Accessed: 22 May 2018

Dixit, R. (2016) *Think big, act small: mobile video advertising in India and Indonesia, Google* [Online]. Available at: https://www.thinkwithgoogle.com/intl/en-apac/tools-research/research-

studies/think-big-act-small-mobile-video-advertising-in-india-and-indonesia (Accessed: 26 April 2017)

Doe, P. and Kugel, C. (2006) *A powerful mix - how people, TV and internet interact* [Online]. Available at: https://ana.esomar.org/documents/a-powerful-mix-6906 (Accessed: 03 January 2019)

Dominick, J.R. (2010) *The dynamics of mass communication: media in the digital age.* New York: McGraw-Hill Education, pp. 11–37.

Domo. (2020) *Media usage in an internet minute as of August 2020* [Online]. Available at: https://www.statista.com/statistics/195140/new-user-generated-content-uploaded-by-users-per-minute (Accessed: 01 February 2021)

Donnenfeld, S. and Goodhand, A. (2000) 'Digital kids—navigating and making sense of a world of choice', *International Journal of Advertising and Marketing to Children*, 1(4), pp. 291–302.

Draper, N.A. and Turow, J. (2019) 'The corporate cultivation of digital resignation', *New Media & Society*, 21(8), pp. 1824–1839.

Drolet, A., Simonson, I. and Tversky, A. (2000) 'Indifference curves that travel with the choice set', *Marketing Letters*, 11(3), p.199.

Du Plessis, E. (1994) 'Understanding and using likability', *Journal of Advertising Research*, 34(5), pp. RC3+.

Du Plessis, E. (2006) *Commap ad likeability model, book of brand advertising models* [Online]. Available at: https://www.warc.com/content/article/swocc/commap_ad_likeability_model/88056 (Accessed: 27 February 2019)

Dwyer, T. (2010) *Media convergence*. UK: McGraw-Hill Education.

Dwyer, T. (2015) *Convergent media and privacy*. London: Springer.

Eagly, A.H. and Chaiken, S. (1998) 'Attitude structure and function' in Gilbert, D.T., Lindzey, G. and Fiske, S.T. (eds.) *Handbook of social psychology, 2*. UK: Wiley, pp. 269–322.

Edmonds, W.A. and Kennedy, T.D. (2012) *An applied reference guide to research designs: quantitative, qualitative, and mixed methods.* London: SAGE Publications.

Ehrenberg, A.S. (2000) 'Repetitive advertising and the consumer', *Journal of Advertising Research*, 40(6), pp. 39–48.

Ehrenberg, A.S. and Charlton, P. (1973) 'An analysis of simulated brand choice', *Journal of Advertising Research*.

Einstein, M. (2017) *Advertising: what everyone needs to know®*. New York: Oxford University Press.

Ekman, P. (1992) 'Are there basic emotions?', *Psychological Review,* 99(3), pp. 550–553. doi: 10.1037/0033-295X.99.3.550

Ellis-Petersen, H. (2020) 'India bans TikTok after Himalayan border clash with Chinese troops', *The Guardian*, 29 June [online]. Available at: https://www.theguardian.com/world/2020/jun/29/india-bans-tiktok-after-himalayan-border-clash-with-chinese-troops (Accessed: 26 November 2020)

Ellul, J. (1973) *Propaganda: the formation of men's attitudes* (pp. 116-120). New York: Vintage Books.

Elmer, G. (2003) *Profiling machines: mapping the personal information economy*. Cambridge, MA: MIT Press.

Ernst & Young (2019) *A billion screens of opportunity: India's media and entertainment sector* [Online]. Available at: https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/tmt/ey-how-a-billion-screens-can-turn-india-into-a-m-and-e-powerhouse.pdf (Accessed: 24 April 2020)

Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K. and Kirlik, A. (2016) First I "like" it, then I hide it: folk theories of social feeds. *Proceedings of the 2016 CHI conference on human factors in computing systems, May 2016,* New York: Association for Computing Machinery, pp. 2371–2382.

Facebook (2019) *Find your audience* [Online]. Available at: https://www.facebook.com/business/ads/ad-targeting (Accessed: 31 January 2019)

Fieldhouse, M. and Nicholas, D. (2008) 'Digital literacy as information savvy: The road to information literacy', *Digital Literacies: Concepts, Policies and Practices*, pp. 47–72.

Findlay, S. (2020) 'India bans TikTok and dozens of other Chinese mobile apps over privacy fears', *Financial Times*, 30 June [online]. Available at: https://www.ft.com/content/08e15c26-48e0-4540-a040-1a8782e84f2e (Accessed: 26 November 2020)

Fiore, Q. and McLuhan, M. (1967) *The medium is the massage*. New York: Penguin Random House.

Flyverbom, M. (2019) *The digital prism: transparency and managed visibilities in a datafied world*. Cambridge: Cambridge University Press.

Fotopoulou, A. (2018) 'From networked to quantified self: self-tracking and the moral economy of data sharing', in Papacharissi, Z. (ed.) *A networked self and platforms, stories, connections.* New York: Routledge, pp. 144–159.

Foucault, M. (2008) '"Panopticism" from "discipline & punish: the birth of the prison"', *Race/Ethnicity: Multidisciplinary Global Contexts*, 2(1), pp. 1–12. http://www.jstor.org/stable/25594995

Fuchs, C. (2012) 'Dallas Smythe today – the audience commodity, the digital labour debate, Marxist political economy and critical theory', *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, 10(2), pp. 692–740.

Fulgoni, G. M. (2018) 'Are you targeting too much? Effective marketing strategies for brands', *Journal of Advertising Research*, 58(1), pp. 8–11. doi: 10.2501/JAR-2018-008.

Ganascia, J. G. (2010) 'The generalized sousveillance society'. *Social Science Information*, 49(3), pp. 489–507.

Gandy Jr, O. H. (1989) 'The surveillance society: information technology and bureaucratic social control', *Journal of Communication*, 39(3), pp. 61–76.

Gandy Jr, O. H. (1993) *The panoptic sort: a political economy of personal information*. New York: Oxford University Press.

Gandy, R. (1996) 'Human versus mechanical intelligence'. *Machines and Thought: The Legacy of Alan Turing*, Oxford: Clarendon, pp. 125–136.

Gehl, R.W. (2015) 'Sharing, knowledge management and big data: a partial genealogy of the data scientist', *European Journal of Cultural Studies*, 18(4-5), pp. 413–428.

Georgakopoulou, A. and Spilioti, T. (eds). (2015) *The Routledge handbook of language and digital communication*. Abingdon: Routledge.

Ghosh, B. and Stock, A. (2010) 'Advertising effectiveness, digital video recorders, and product market competition', *Marketing Science*, 29(4), pp. 639–649. doi: 10.1287/mksc.1090.0544.

Gillespie, T. (2014) 'The relevance of algorithms' in Gillespie, T., Boczkowski, P. J. and Foot, K. A. (eds.) *Media technologies: essays on communication, materiality, and society*. Cambridge, MA: MIT Press, p. 167.

Gilliom, J. (2001) *Overseers of the poor: surveillance, resistance, and the limits of privacy*. Chicago: University of Chicago Press.

Glass, R. and Callahan, S. (2014) *The big data-driven business: how to use big data to win customers, beat competitors, and boost profits*. Hoboken, New Jersey: John Wiley & Sons.

Golafshani, N. (2003) 'Understanding reliability and validity in qualitative research', *The Qualitative Report*, 8(4), pp. 597–607.

Google, 2017 consumer barometer [Online]. Available at: https://www.consumerbarometer.com/en/graphbuilder/?question=M1&filter=country:india,united_kingdom,united_states (Accessed: 24 April 2018)

Goyal, M. (2020) '25 years of internet in India and how it has transformed our lives',' *The Economic Times,* 16 August [online]. Available at: https://economictimes.indiatimes.com/tech/internet/the-india-story-after-25-years-of-internet-and-how-it-has-transformed-our-lives/articleshow/77565596.cms (Accessed: 22 October 2020)

Granovetter, M., 1983. The strength of weak ties: A network theory revisited [Online]. Available at: https://www.cs.kent.ac.uk/people/staff/saf/share/great-missenden/reference-papers/Weak%20Ties.pdf (Accessed: 27 May 2017)

Green, A. (2017) *Audience measurement 5.0 – pushing the boundaries* [Online]. Available at: https://www.ipsos.com/en/audience-measurement-50-pushing-boundaries (Accessed: 23 April 2019)

Grimshaw, C. (2017) *A post-demographic age*. [Online]. Available at: https://www.warc.com/content/article/a-post-demographic-age/en-gb/117252 (Accessed: 27 May 2018)

Gurses, S., van Hoboken, J. V. J. (2017) *Privacy after the agile turn* [Online]. Available at: https://osf.io/preprints/socarxiv/9gy73/ (Accessed: 23 April 2021)

Ha, L. (2008) 'Online advertising research in advertising journals: a review,' *Journal of Current Issues & Research in Advertising*, 30(1), pp. 31–48.

Haggin, P. and Vranica, S. (2021) 'Apple's moves vex digital advertisers', *The Wall Street Journal* (English edition), 09 June [online]. Available at: https://www.wsj.com/articles/apples-moves-to-tighten-flow-of-user-data-leave-advertisers-anxious-11623189903 (Accessed: 26 June 2021)

Hanssens, D.M. (2009) 'Advertising impact generalizations in a marketing mix context', *Journal of Advertising Research*, 49(2), pp. 127–129.

Hartman, T., Kennedy, H., Steedman, R. and Jones, R. (2020) 'Public perceptions of good data management: findings from a UK-based survey', *Big Data and Society*, 7(1). doi: 10.1177/2053951720935616

Hassoun, D. (2014) 'Tracing attentions: toward an analysis of simultaneous media use', *Television & New Media*, 15(4), pp. 271–288.

Heo, J. and Cho, C.H. (2009) 'A new approach to target segmentation: media-usage segmentation in the multi-media environment', *Journal of Targeting, Measurement and Analysis for Marketing*, 17(3), pp. 145–155.

Hepp, A., Breiter, A. and Friemel, T. N. (2018) 'Digital traces in context | digital traces in context—an introduction', *International Journal of Communication*, 12, p.11.

Herman, E. and Chomsky, N. (2012) '17 A propaganda model' in Durham, M.G. and Kellner, D.M. (eds.) *Media and cultural studies: keyworks.* Oxford: Wiley-Blackwell, p.204.

Heukelom, F. (2012) 'Three explanations for the Kahneman-Tversky programme of the 1970s', *European Journal of the History of Economic Thought*, 19(5), pp. 797–828.

Hosanagar, K. (2020) *A human's guide to machine intelligence: how algorithms are shaping our lives and how we can stay in control.* London: Penguin Books.

Humphreys, L. (2011) 'Who's watching whom? a study of interactive technology and surveillance', *Journal of Communication,* 61(4), pp. 575–595.

IAMAI and Kantar IMRB (2017) *Internet in India 2017* [Online]. Available at: http://www.mxmindia.com/wp-content/uploads/2018/02/Internet-in-india_10-1-18.compressed.pdf (Accessed: 27 October 2018)

Iliadis, A. and Russo, F. (2016) 'Critical data studies: an introduction', *Big Data & Society*, 3(2), p. 2053951716674238.

Indian Newspaper Society (2003) *Impact Multiplier* [Online]. Available at: https://www.afaqs.com/news/story/6997_IMRBs-Impact-Multiplier-study-to-be-released-on-September-8 (Accessed: 31 March 2019)

Indian Parliament (2019) *Personal data protection bill 2019* [Online]. Available at: http://164.100.47.4/BillsTexts/LSBillTexts/Asintroduced/373_2019_LS_Eng.pdf (Accessed: 23 April 2021)

Izard, C. E. (1992) 'Basic emotions, relations among emotions, and emotion-cognition relations,' *Psychological Review, 99*(3), pp. 561–565. doi: 10.1037/0033-295X.99.3.561

Jain, S. (2015) 'India: multiple media explosions', in Nordenstreng, K. and Thussu, D.K. (eds.) *Mapping BRICS media.* New York: Routledge, pp. 145–165.

James, C. (2009) *Young people, ethics, and the new digital media: a synthesis from the GoodPlay project*. Cambridge, MA: MIT Press, p.128.

Järvinen, J., Tollinen, A., Karjaluoto, H. and Jayawardhena, C. (2012) 'Digital and social media marketing usage in b2b industrial section', *Marketing Management Journal*, 22(2), pp. 102–117.

Jenkins, H. (2006) *Convergence culture*. New York: New York University Press.

Jenkins, H. (2006) Who the &% &# Is Henry Jenkins? c*onfessions of an ACA-fan* [Online]. Available at: http://henryjenkins.org/blog/2006/06/who_the_is_henry_jenkins.html (Accessed: 27 May 2017)

Johnson, G. A., Lewis, R. A. and Nubbemeyer, E. I. (2017) 'Ghost ads: improving the economics of measuring online ad effectiveness', *Journal of Marketing Research (JMR*), 54(6), pp. 867–884. doi: 10.1509/jmr.15.0297.

Johnson, R. B. (1997) 'Examining the validity structure of qualitative research', *Education,* 118(2), pp. 282–293.

Jowett, G. S. and O'donnell, V. (2018) *Propaganda & Persuasion*. California: SAGE Publications.

Jung, Y., Perez-Mira, B., and Wiley-Patton, S. (2009) 'Consumer adoption of mobile TV: examining psychological flow and media content', *Computers in Human Behavior*, 25(1) pp. 123–129.

Kahneman, D. (2003) 'Maps of bounded rationality: psychology for behavioral economics', *American Economic Review*, 93(5), pp. 1449–1475.

Kahneman, D. and Snell, J. (1992) 'Predicting a changing taste: do people know what they will like?', *Journal of Behavioral Decision Making*, 5(3), pp. 187–200.

Kahneman, D., Lovallo, D. and Sibony, O. (2011) 'Before you make that big decision..', *Harvard Business Review*, 89(6), pp. 50–60.

Kandias, M., Mitrou, L., Stavrou, V. and Gritzalis, D. (2013) *Which side are you on? a new panopticon vs. privacy* [Online]. Available at: https://ieeexplore.ieee.org/document/7223159 (Accessed: 23 January 2019)

Kantar IMRB (2019) *21st edition ICUBE digital adoption & usage trends* [Online]. Available at: https://imrbint.com/images/common/ICUBE%E2%84%A2_2019_Highlights.pdf (Accessed: 6 September 2021)

Kantchev, G. and Gross, J. (2018) 'Cambridge university pushes back against Facebook scrutiny of research apps', *The Wall Street Journal* (English edition), 14 April [online]. Available at: https://www.wsj.com/articles/cambridge-university-pushes-back-against-facebook-scrutiny-of-research-apps-1523703601 (Accessed: 26 March 2020)

Kaplan, M. (2020) 'The self-consuming commodity: audiences, users, and the riddle of digital labor', *Television & New Media*, 21(3), pp. 240–259.

Katz, E., Lazarsfeld, P.F. and Roper, E. (1966) *Personal influence: the part played by people in the flow of mass communications*. Toronto, Ontario: The Free Press

Kennedy, H. (2016) *Post, mine, repeat: social media data mining becomes ordinary*. London: Springer.

Kennedy, H. (2018) 'Living with data: aligning data studies and data activism through a focus on everyday experiences of datafication', *Krisis: Journal for Contemporary Philosophy,* 2018(1), pp. 18–30.

Kennedy, H. and Bates, J. (2017) 'Data power in material contexts: Introduction', *Television & New Media*, 18(8), pp. 701–705.

Kennedy, H. and Hill, R.L. (2017) 'The pleasure and pain of visualizing data in times of data power', *Television & New Media*, 18(8), pp. 769–782.

Kerr, G., Schultz, D.E., Kitchen, P.J., Mulhern, F.J. and Beede, P. (2015) 'Does traditional advertising theory apply to the digital world? : a replication analysis questions the relevance of the elaboration likelihood model', *Journal of Advertising Research*, 55(4), pp. 390–400.

Kim, K.J. and Sundar, S.S. (2016) 'Mobile persuasion: can screen size and presentation mode make a difference to trust?', *Human Communication Research*, 42(1), pp. 45–70.

Kitchin, R. (2014) *The data revolution: big data, open data, data infrastructures and their consequences*. London: SAGE Publications.

Kitchin, R. and Dodge, M. (2014) *Code/space: software and everyday life*. Cambridge, MA: MIT Press.

Klein, G. A. (1999) *Sources of power: how people make decisions*. Cambridge, MA: MIT Press.

Klosowski, T. (2021) 'The state of consumer data privacy laws in the US (and why it matters)', *The New York Times Wirecutter,* 06 September [online]. Available at: https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us (Accessed: 09 September 2021)

Ko, H., Cho, C.H. and Roberts, M.S. (2005) 'Internet uses and gratifications: a structural equation model of interactive advertising', *Journal of Advertising*, 34(2), pp. 57–70.

Kohli-Khandekar, V. (2013) *The Indian media business*. New Delhi: SAGE Publications India.

Krugman, H.E. (1965) 'The impact of television advertising: learning without involvement', *Public Opinion Quarterly*, 29(3), pp. 349–356.

Lafley, A. G. and Martin, R. L. (2017) 'Customer loyalty is overrated (cover story)', *Harvard Business Review*, 95(1), pp. 45–54.

Larkin, J. (2005*)* 'Rural India goes digital; satellite TV pushes into remote areas, meeting pent-up demand', *The Wall Street Journal* (English edition)*,* 03 October [online]. Available at: https://www.wsj.com/articles/SB112829847867958098 (Accessed: 26 June 2019)

Lasswell Harold, D. (1927*) Propaganda technique in the world war*. New York: AA Knopf.

Lasswell, H.D. (1948) 'The structure and function of communication in society', *The Communication of Ideas*, 37, pp. 215–228.

Lazarsfeld, P.F., Berelson, B. and Gaudet, H. (1948) *The people's choice: how the voter makes up his mind in a presidential campaign.* 2nd edn. New York: Columbia University Press.

Le, T.D. and Nguyen, B.T.H. (2014) 'Attitudes toward mobile advertising: a study of mobile web display and mobile app display advertising', *Asian Academy of Management Journal*, 19(2), p.87.

Lecinski, J. (2011) *Winning the zero moment of truth* [Online]. Available at: https://www.thinkwithgoogle.com/marketing-resources/micro-moments/2011-winning-zmot-ebook/ (Accessed: 16 December 2018)

Leckenby, J.D. and Kim, H. (1994) 'How media directors view reach/frequency estimation: now and a decade ago', *Journal of Advertising Research*, 34(5), pp. 9–22.

Lederer, C. and Hill, S. (2001) 'See your brands through your customers' eyes', *Harvard Business Review, 79(6),* pp. 125–133.

Lee, M. (2011) 'Google ads and the blindspot debate', *Media, Culture & Society*, 33(3), pp. 433–447.

Levy, S. (2021) *In the plex: how Google thinks, works, and shapes our lives.* New York: Simon & Schuster.

Lim, J.S., Ri, S.Y., Egan, B.D. and Biocca, F.A. (2015) 'The cross-platform synergies of digital video advertising: implications for cross-media campaigns in television, internet and mobile TV', *Computers in Human Behavior*, 48, pp. 463–472.

Lin, C., Venkataraman, S. and Jap, S. D. (2013) 'Media multiplexing behavior: Implications for targeting and media planning', *Marketing Science*, 32(2), pp. 310–324.

Lin, C.A. (1994) 'Audience fragmentation in a competitive video marketplace', *Journal of Advertising Research*, 34(6), pp. 30–39.

Lomborg, S. and Kapsch, P. H. (2020) 'Decoding algorithms', *Media, Culture & Society,* 42(5), pp. 745–761.

Lumen Research (2018) *Welcome to the attention economy: How to align your media buying strategy with consumer attention for the higher brand recall and ROI.* [Online]. Available at: https://dkf1ato8y5dsg.cloudfront.net/uploads/8/53/cpn-avocet-whitepaper-final.pdf (Accessed: 23 April 2019)

Lupton, D. (2016) *The quantified self*. Cambridge: Polity.

Lupton, D. (2018) 'How do data come to matter? Living and becoming with personal data'. *Big Data & Society*, 5(2), doi: 10.1177/2053951718786314.

Lupton, D. (2020a) 'Better understanding about what's going on': young Australians' use of digital technologies for health and fitness', *Sport, Education and Society*, 25(1), pp. 1–13.

Lupton, D. (2020b) *Data selves: more-than-human perspectives*. London: Polity.

Lupton, D. (2021) ''Not the real me': social imaginaries of personal data profiling', *Cultural Sociology,* 15(1), pp. 3–21.

Lyon, D. (2014) 'Surveillance, Snowden, and big data: capacities, consequences, critique', *Big Data & Society*, 1(2), p.2053951714541861.

Mahnke, M. and Uprichard, E. (2014) 'Algorithming the algorithm' in Rasch, M. and König, R. (eds.) *Society of the query reader: reflections on web search.* Amsterdam: Institute of Network Cultures, pp. 256–270.

Martin, L. (2019) 'PewDiePie vs T-Series update: YouTube star makes heartfelt plea to subscribers', *Express*, 02 May [online]. Available at: https://www.express.co.uk/entertainment/gaming/1120884/PewDiePie-vs-T-Series-meme-subscribers-sub-count-Christchurch (Accessed: 26 July 2019)

Masih, N. (2020) *India's TikTok ban eats into his empire* [Online]. Available at: https://www.pressreader.com/usa/the-washington-post/20200711/281956020080738 (Accessed: 23 September 2020)

Mathiesen, T. (1997) 'The viewer society: Michel Foucault's panopticon' revisited'. *Theoretical criminology*, 1(2), pp. 215–234.

Mayernik, M. S. and Acker, A. (2018) 'Tracing the traces: the critical role of metadata within networked communications', *Journal of the Association for Information Science and Technology*, 69(1), pp. 177–180.

Mayer-Schönberger, V. (2009) *Delete: the virtue of forgetting in the digital age. E-book library* [online]. Available at: https://www.proquest.com/docview/1681853783 (Accessed: 20 October 2020)

Mayer-Schönberger, V. and Cukier, K. (2013) *Big data: a revolution that will transform how we live, work, and think*. New York: Houghton Mifflin Harcourt.

McCombs, M.E. and Shaw, D.L. (1993) 'The evolution of agenda-setting research: twenty-five years in the marketplace of ideas', *Journal of Communication*, 43(2), pp. 58–67.

McDonald, J. and Clapp, R. (2020) *Global ad trends: state of the industry 2020/21* [Online]. Available at: https://www.warc.com/content/article/warc-data/global-ad-trends-outlook-2020-21/134804 (Accessed: 23 June 2020)

McGuire, W. (1999) *Constructing social psychology: creative and critical aspects*. Cambridge: Cambridge University Press.

McKinnon, J. D. (2020) 'TikTok shutdown argued before federal judge', *The Wall Street Journal* (English edition), 04 November [online]. Available at: https://www.wsj.com/articles/tiktok-shutdown-argued-before-federal-judge-11604518470 (Accessed: 20 January 2021)

McQuail, D. (1997) *Audience analysis*. London: SAGE Publications.

McStay, A. (2009) *Digital advertising*. New York: Palgrave Macmillan.

McStay, A. (2011) 'Profiling phorm: an autopoietic approach to the audience-as-commodity', *Surveillance & Society*, 8(3), pp. 310–322.

McStay, A. (2017) 3.1 Micro-Moments, liquidity, intimacy and automation: developments in programmatic ad-tech in Siegert, G., Rimscha, M. and Grubenmann, S. (eds.) *Commercial communication in the digital age: information or disinformation?* Berlin, Boston: De Gruyter Saur, pp. 143–160. doi: 10.1515/9783110416794-009

McStay, A. (2018) 'Digital advertising and adtech: programmatic platforms, identity and moments', in Powell, H., MacRury, I. and Hardy, J. (eds.) *The advertising handbook.* New York: Routledge, pp. 88–100.

Meehan, E. R. (1984) 'Ratings and the institutional approach: a third answer to the commodity question', *Critical Studies in Mass Communication*, 1(2), pp. 216–25.

Meehan, E. R. and Torre, P. J. (2011) 'Markets in theory and markets in television' in Wasko, J., Murdock, G. and Sousa, H. (eds). *The handbook of political economy of communications*. Oxford: Wiley Blackwell, pp. 62-82. doi: 10.1002/9781444395402.ch3

Meehan, E.R. (2007) 'Understanding how the popular becomes popular: The role of political economy in the study of popular communication', *Popular Communication*, 5(3), pp. 161–170.

Meehan, E.R. (2012) 'Gendering the commodity audience: critical media research, feminism, and political economy' in Durham, M.G. and Kellner, D.M. (eds.) *Media and cultural studies: keyworks.* Oxford: Wiley-Blackwell, pp. 242–249.

Meeker, M. (2019) *Internet trends 2019* [Online]. Available at: https://www.bondcap.com/report/itr19/ (Accessed: 01 February 2021)

Meyers, C.B. (1997) 'Frank and Anne Hummert's soap opera empire: "reason-why" advertising strategies in early radio programming', *Quarterly Review of Film & Video,* 16(2), pp. 113–132.

Meyers-Levy, J. and Malaviya, P. (1999) 'Consumers' processing of persuasive advertisements: an integrative framework of persuasion theories', *Journal of Marketing*, 63(4_suppl1), pp. 45–60.

Meyrowitz, J. (2009) 'We liked to watch: television as progenitor of the surveillance society'. *The Annals of the American Academy of Political and Social Science*, 625(1), pp. 32–48.

Michael Lithgow, M. (2021) *Pedagogies of the datafied: material foundations for literacies of the subject in the 21st century* [Online]. Available at: https://static1.squarespace.com/static/5cf15af7a259990001706378/t/60b69866e1637f34f9fec0fa/1622579303985/Lithgow+Pedagogies+of+the+Datafied.pdf (Accessed: 24 July 2021)

Moran, G., Muzellec, L. and Nolan, E. (2014) 'Consumer moments of truth in the digital context: how "search" and "e-word of mouth" can fuel consumer decision making', *Journal of Advertising Research*, 54(2), pp. 200–204.

Naik, P.A. and Peters, K. (2009) 'A hierarchical marketing communications model of online and offline media synergies', *Journal of Interactive Marketing*, 23(4), pp. 288–299.

Naik, P.A. and Raman, K. (2003) 'Understanding the impact of synergy in multimedia communications', *Journal of Marketing Research*, 40(4), pp. 375–388.

Napoli, P. M. (2014) 'Automated media: an institutional theory perspective on algorithmic media production and consumption', *Communication Theory*, 24(3), pp. 340–360.

Napoli, P.M. (2001) 'The audience product and the new media environment: implications for the economics of media industries', *International Journal on Media Management*, 3(2), pp. 66–73.

Napoli, P.M. (2018) *What if more speech is no longer the solution: first amendment theory meets fake news and the filter bubble* [Online]. Available at: https://www.thefreelibrary.com/What+If+More+Speech+Is+No+Longer+the+Solution%3f+First+Amendment+Theory...-a0539774158 (Accessed: 23 April 2019)

Nelson, E. and Ellison, S. (2005) 'In a shift, marketers beef up ad spending inside stores; funky displays and lighting, TV spots in Wal-Mart; unsettling Madison Avenue', *The Wall Street Journal* (English edition), 21 September [online]. Available at: https://www.wsj.com/articles/SB112725891535046751 (Accessed: 26 June 2019)

Nelson, K.A. (2004) 'Consumer decision making and image theory: understanding value-laden decisions', *Journal of Consumer Psychology (Lawrence Erlbaum Associates)*, 14(1/2), pp. 28–40.

Neustar, I. (2018) *Neustar media effectiveness study reveals digital advertising tops TV in movie studio mix* [Online]. Available at: https://www.home.neustar/about-us/news-room/press-releases/2018/MediaEffectivenessStudios (Accessed: 19 December 2018).

Nguyen, P. and Solomon, L. (2018) *Consumer data and the digital economy* [Online]. Available at: https://cprc.org.au/app/uploads/2018/07/Consumer-Data-and-the-Digital-Economy_smallest-file-size.pdf (Accessed: 20 February 2021)

Nieborg, D.B. (2017) 'App advertising: the rise of the player commodity' in Korin, E. and Hamilton, J.F. (eds.) *Explorations in critical studies of advertising*. New York: Routledge, pp. 28–41.

Nielsen, R. K. (2016) 'Folk theories of journalism: the many faces of a local newspaper', *Journalism Studies*, 17(7), pp. 840–848.

Nightingale, V. (ed.) (2013) *The handbook of media audiences*. Oxford: Wiley-Blackwell.

Nikolaeva, E. E., and Berendeeva, A. B. (2020) *1. Man as an element of the economic mechanism of the information society in modern social and economic theories and concepts* [Online]. Available at: https://doi.org/10.1515/9783110636147-001 (Accessed 21 October 2020)

Ninan, T. N. (2018) 'The political economy of reforms: the art of the possible' in Mohan, R. (ed.) *India transformed: twenty-five years of economic reforms.* Washington, D.C.: Brookings Institution Press, pp. 77–90.

Nord VPN (2020) *National privacy test* [Online]. Available at: https://s1.nordcdn.com/nord/misc/0.34.0/vpn/brand/national-privacy-test-report.pdf (Accessed: 06 September 2021)

Nordenstreng, K. and Thussu, D. K. (eds.) (2015) *Mapping BRICS media*. New York: Routledge.

Novemsky, N. and Kahneman, D. (2005) 'The boundaries of loss aversion', *Journal of Marketing Research (JMR)*, 42(2), pp. 119–128.

O. Nyumba, T., Wilson, K., Derrick, C. J. and Mukherjee, N. (2018) 'The use of focus group discussion methodology: insights from two decades of application in conservation', *Methods in Ecology and Evolution*, 9(1), pp. 20–32.

O'Neil, C. (2016*) Weapons of math destruction*. London: Penguin Books.

Osaki, T. (2008) 'Global marketing strategy in the digital age: an analysis of Japanese mobile phone manufacturers', *Marketing Review*, 8(4), pp. 329–341.

Palmer, J. (1992) 'Women and soap opera', *Theory, Culture & Society*, 9(3), pp. 125–138.

Palmer, R. A. (2019) 'A "deep story" about American journalism: using "episodes" to explore folk theories of journalism', *Journalism Studies*, 20(3), pp. 327–344.

Pangrazio, L. and Selwyn, N. (2019) ''Personal data literacies': a critical literacies approach to enhancing understandings of personal digital data', *New Media & Society*, 21(2), pp. 419–437.

Papacharissi, Z. (ed.) (2010) *A networked self: identity, community, and culture on social network sites*. New York: Routledge.

Pare, V. and Dawes, J. (2007) *Deviations from double jeopardy: how many private label and high-share brands exhibit excess loyalty?* ANZMAC, Dunedin, New Zealand, 2007, pp. 1717–1722.

Park, M. J. and Curran, J. (eds.) (2000) *De-Westernizing media studies*. New York: Routledge.

Pasquale, F. (2015) *The black box society: the secret algorithms that control money and information.* Cambridge, MA and London, England: Harvard University Press. https://doi.org/10.4159/harvard.9780674736061

Paul, S. L. and Wang, G. (1995) 'Satellite TV in Asia: forming a new ecology', *Telecommunications Policy*, 19(2), pp. 135–149.

Peacock, S.E. (2014) 'How web tracking changes user agency in the age of big data: The used user', *Big Data & Society*, 1(2), p.2053951714564228.

Petty, R.E. and Cacioppo, J.T. (1986) 'The elaboration likelihood model of persuasion' in Petty, R.E. and Cacioppo, J.T. (eds.) *Communication and persuasion*. New York: Springer, pp. 1–24.

Petty, R.E. and Wegener, D.T. (1998) 'Attitude change: multiple roles for persuasion variables' in Fiske, S.T., Gilbert, D.T. and Lindzey, G. (eds.) *The handbook of social psychology, 1*. UK: Wiley, pp. 323–390

Petty, R.E., Cacioppo, J.T. and Schumann, D. (1983) 'Central and peripheral routes to advertising effectiveness: the moderating role of involvement', *Journal of Consumer Research*, 10(2), pp. 135–146.

Pilotta, J, Schultz, D, Drenik, G, & Rist, P. (2004) 'Simultaneous media usage: a critical consumer orientation to media planning', *Journal of Consumer Behaviour,* 3 (3) pp. 285–292.

Pine, B. J. and Gilmore, J. H. (1998) 'Welcome to the experience economy', *Harvard Business Review*, 76, pp. 97–105.

Phillips, D. J. (2004) 'Privacy policy and PETs: The influence of policy regimes on the development and social implications of privacy enhancing technologies', *New Media & Society*, 6(6), pp. 691–706. https://doi-org.soton.idm.oclc.org/10.1177/146144804042523.

Pink, S., Postill, J., Mackley, K.L. and Astari, N. (2017) 'Digital-visual stakeholder ethnography', *Sociological Research Online*, (4), pp. 174–192. doi: 10.1177/1360780417726736.

Pirolli, P. (2007) *Information foraging theory: adaptive interaction with information*. *E-book library* [online]. Available at: https://oxford-universitypressscholarship-com.soton.idm.oclc.org/view/10.1093/acprof:oso/9780195173321.001.0001/acprof-9780195173321 (Accessed: 19 January 2021)

Prensky, M. (2001) Digital natives, digital immigrants [Online]. Available at: https://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1 (Accessed: 27 May 2017)

Protess, D. and McCombs, M. E. (2016) *Agenda setting: readings on media, public opinion, and policymaking*. New York: Routledge.

Purnell, N. (2017) 'Cellphone bills in India inch closer to zero', *The Wall Street Journal* (English edition), 21 July [online]. Available at: https://www.wsj.com/articles/reliances-ambani-launches-low-cost-mobile-phone-for-india-1500629140 (Accessed: 31 March 2019)

Pybus, J., Coté, M. and Blanke, T. (2015) 'Hacking the social life of big data', *Big Data & Society*, 2(2), p.2053951715616649.

Quinton, S. (2013) 'The community brand paradigm: a response to brand management's dilemma in the digital era', *Journal of Marketing Management*, 29(7-8), pp. 912–932.

Rabiee, F. (2004) 'Focus-group interview and data analysis', *Proceedings of The Nutrition Society*, 63(4), pp. 655–660.

Raley, R. (2013) 'Dataveillance and countervailance' in Gitelman, L. (ed.) *'Raw data' is an oxymoron*. Cambridge, MA: MIT Press, pp. 121–146.

Reynolds, T. J. (2006) 'Methodological and strategy development implications of decision segmentation', *Journal of Advertising Research*, 46(4), pp. 445–461.

Rieder, B. and Sire, G. (2014) 'Conflicts of interest and incentives to bias: A microeconomic critique of Google's tangled position on the web', *New Media & Society*, 16(2), pp. 195–211.

Rip, A. (2006) 'Folk theories of nanotechnologists', *Science as Culture*, 15(4), pp. 349–365.

Rodgers, S., Ouyang, S. and Thorson, E. (2017) 'Revisiting the interactive advertising model (IAM) after 15 years', *Digital Advertising: Theory and Research*, pp. 1–18.

Rogers, E. M. (1995) *Diffusion of innovations*. New York: Free Press.

Rogers, E.M. and Shoemaker, F.F. (1971*) Communication of innovations; a cross-cultural approach.* New York: Free Press.

Rogers, R. (2009) *Post-demographic machines: studying social networking* [Online]. Available at: http://www.govcom.org/publications/full_list/WalledGarden_ch04_RR.pdf (Accessed: 25 November 2018)

Rossiter, J. R. and Danaher, P. J. (1998) *Advanced media planning Vol. (1).* New York: Springer.

Roy, R. and Li, S. (2020) 'India bans TikTok, dozens of other Chinese apps after border clash; UC Browser and WeChat are among the banned mobile apps*', The Wall Street Journal* (English edition)*,* 30 June [online]. Available at: https://www.wsj.com/articles/india-blocks-dozens-of-chinese-apps-including-tiktok-following-border-clash-11593447321 (Accessed: 15 July 2020)

Rubinson, J. (2009) 'Empirical evidence of TV advertising effectiveness', *Journal of Advertising Research*, 49(2), pp. 220–226.

Ruckenstein, M. and Granroth, J. (2020) 'Algorithms, advertising and the intimacy of surveillance', *Journal of Cultural Economy,* 13(1), pp. 12–24.

Rungie, C. and Goodhardt, G. (2004) 'Calculation of theoretical brand performance measures from the parameters of the Dirichlet model', *Marketing Bulletin*, 15, pp. 1–19.

S.Wall, D. (2019) 'Surveillant internet technologies and the growth in information capitalism: spams and public trust in the information society' in Haggerty, K. and Ericson, R. (eds.) *The new politics of surveillance and visibility*. Toronto: University of Toronto Press, pp. 340–362. doi: 10.3138/9781442681880-015

Sadowski, J. (2020) *Too smart: how digital capitalism is extracting data, controlling our lives, and taking over the world.* Cambridge, MA: MIT Press.

Sandel, M. J. (2012) *What money can't buy: The moral limits of markets.* New York: Farrar, Straus and Giroux.

SanJosé-Cabezudo, R., Gutiérrez-Arranz, A.M. and Gutiérrez-Cillán, J. (2009) 'The combined influence of central and peripheral routes in the online persuasion process', *CyberPsychology & Behavior*, 12(3), pp. 299–308.

Saulles, M. D. and Horner, D. S. (2011) 'The portable panopticon: morality and mobile technologies', *Journal of Information, Communication & Ethics in Society*, 9(3), pp. 206–216.

Schechner, S., Grind, K. and Haggin, P. (2020) 'Tech firms are spying on you. In a pandemic, governments say that's OK', *The Wall Street Journal* (English edition), 15 June [online]. Available at: https://www.wsj.com/articles/once-pariahs-location-tracking-firms-pitch-themselves-as-covid-sleuths-11592236894 (Accessed: 11 July 2020)

Sharp, B. and Wind, Y. (2009) 'Today's advertising laws: will they survive the digital revolution?', *Journal of Advertising Research*, 49(2), pp. 120–126.

Shklovski, I., Mainwaring, S.D., Skúladóttir, H.H. and Borgthorsson, H. Leakiness and creepiness in app space: perceptions of privacy and mobile app use. *Proceedings of the SIGCHI conference on human factors in computing systems, April 2018,* Montreal: pp. 2347–2356.

Shoenberger, H. (2017) 'Targeted digital advertising and privacy' in Rodgers, S. and Thorson, E. (eds.) *Digital advertising: theory and research.* New York: Routledge, pp. 300–309.

Shrum, L.J., Liu, M., Nespoli, M. and Lowrey, T.M. (2013) 'Persuasion in the marketplace : how theories of persuasion apply to marketing and advertising' in Dillard, J.P. and Shen, L. (eds.) *The SAGE handbook of persuasion: developments in theory and practice*. 2nd edn. London: SAGE Publications, pp. 314-330.

Simon, M.R., Van Den Driest, F. and Wilms, T. (2016) 'Driving customer-centric growth: a practical roadmap: the pivotal role of insights and analytics in the customer-centric organization', *Journal of Advertising Research*, 56(2), pp. 159–168.

Simonson, I. and Tversky, A. (1992) 'Choice in context: tradeoff contrast and extremeness aversion', *Journal of Marketing Research (JMR)*, 29, 3, pp. 281–295.

Singh, M. (2019) *Xiaomi tops Indian smartphone market for eighth straight quarter* [Online]. Available at: https://techcrunch.com/2019/08/13/xiaomi-mi-samsung-india-smartphone-market (Accessed: 27 May 2020)

Sissors, J. Z. and Goodrich W. B. (1996) *Media planning workbook*. Lincolnwood: NTC/Contemporary Publishing Group.

Skeggs, B. and Yuill, S. (2016a) 'Capital experimentation with person/a formation: how Facebook's monetization refigures the relationship between property, personhood and protest', *Information, Communication & Society,* 19(3), pp. 380–396.

Skeggs, B. and Yuill, S. (2016b) 'The methodology of a multi-model project examining how Facebook infrastructures social relations', *Information, Communication & Society*, 19(10), pp. 1356–1372.

Smith, E. R. (1998) 'Mental representation and memory', in Gilbert, D. T., Fiske, S. T. and Lindzey, G. (eds.) *The handbook of social psychology, Vol. 1*, 4th edn. New York: McGraw-Hill, pp. 391–445.

Smith, K.T. (2011) 'Digital marketing strategies that millennials find appealing, motivating, or just annoying', *Journal of Strategic Marketing*, 19(6), pp. 489–499.

Smythe, D. W. (1981) 'On the audience commodity and its work' in Durham, M. G and Kellner, D. M. (eds.) *Media and cultural studies: keyworks*. Oxford: Wiley-Blackwell, pp. 230–256.

Solove, D. (2004) *The digital person: technology and privacy in the information age*. New York: New York University Press.

Spindle, B. (2019) ''The youth bulge': India struggles to employ an exploding population', *The Wall Street Journal* (English edition), 22 April [online]. Available at: https://www.wsj.com/articles/the-youth-bulge-india-struggles-to-employ-an-exploding-population-11555925401 (Accessed: 26 August 2020)

Špiranec, S., Kos, D. and George, M. (2019) *Searching for critical dimensions in data literacy* [Online]. Available at: http://informationr.net/ir/24-4/colis/colis1922.html (Accessed: 23 August 2019)

Srinivasan, S., Rutz, O. J. and Pauwels, K. (2016) 'Paths to and off purchase: quantifying the impact of traditional marketing and online consumer activity', *Journal of the Academy of Marketing Science*, 44(4), pp. 440–453.

Steedman, R., Kennedy, H. and Jones, R. (2020) 'Complex ecologies of trust in data practices and data-driven systems', *Information, Communication & Society,* 23(6), pp. 817–832.

Steele, A., Jacobs, D., Siefert, C., Rule, R., Levine, B. and Marci, C.D. (2013) 'Leveraging synergy and emotion in a multi-platform world: a neuroscience-informed model of engagement', *Journal of Advertising Research*, 53(4), pp. 417–430.

Stokes, J. (2012) *How to do media and cultural studies*. London: SAGE Publications.

Straubhaar, J., LaRose, R. and Davenport, L. (2013) *Media now: understanding media, culture, and technology*. Boston, MA: Wadsworth Cengage Learning.

Striphas, T. (2015) 'Algorithmic culture', *European Journal of Cultural Studies*, 18(4-5), pp. 395–412.

Swidler, A. (1986) 'Culture in action: symbols and strategies', *American Sociological Review*, pp. 273–286.

Swisher, K. (2019) 'We're not going to take it anymore', *The New York Times,* 10 April [online]. Available at: https://www.nytimes.com/2019/04/10/opinion/internet-privacy-regulation.html (Accessed: 26 March 2020)

Szulc, L. (2018) 'Profiles, identities, data: making abundant and anchored selves in a platform society', *Communication Theory*, 29(3), pp. 169–188.

Tau, B. (2020) 'Next step in government data tracking is the internet of things', *The Wall Street Journal* (English edition)*,* 27 November [online]. Available at: https://www.wsj.com/articles/next-step-in-government-data-tracking-is-the-internet-of-things-11606478401 (Accessed: 02 January 2021)

Tevi, A. and Koslow, S. (2018) 'How rhetoric theory informs the creative advertising development process: reconciling differences between advertising scholarship and practice', *Journal of Advertising Research*, 58(1), pp. 111–128. doi: 10.2501/JAR-2018-012.

Tewksbury, D. (2005) 'The seeds of audience fragmentation: specialization in the use of online news sites', *Journal of Broadcasting & Electronic Media*, 49(3), pp. 332–348.

Tewksbury, D., 2005. The seeds of audience fragmentation: specialization in the use of online news sites. *Journal of broadcasting & electronic media*, *49*(3), pp.332-348.

Tham, S.M., Rodgers, S. and Thorson, E. (2017) 'Trends and opportunities for digital advertising research' in Rogers, S. and Thorson, E. (eds.) *Digital advertising: theory and research.* New York: Routledge, pp. 31–44.

The Great Hack (2019) Directed by Jehane Noujaim and Karim Ame [Film]. California: Netflix.

The Nielsen Company (2018) *The Nielsen CMO Report 2018* [Online]. Available at: https://www.nielsen.com/us/en/insights/reports/2018/cmo-report-2018-digital-media-roi-measurement-omnichannel-marketing-technology.html (Accessed: 10 March 2019)

The Wall Street Journal (2021) *Inside TikTok's highly secretive algorithm*. 21 July [online]. Available at: https://www.wsj.com/video/series/inside-tiktoks-highly-secretive-algorithm/investigation-how-tiktok-algorithm-figures-out-your-deepest-desires/6C0C2040-FF25-4827-8528-2BD6612E3796 (Accessed: 27 July 2021)

Thomas, J. W. (2017) 'Dark energy in the digital age: Is digital media less effective than traditional media?', *HCM Sales, Marketing & Alliance Excellence Essentials*, 16(3), pp. 9–10.

Thompson, D. (2010) *Google's CEO: 'The laws are written by lobbyists'* Available at: https://www.theatlantic.com/technology/archive/2010/10/googles-ceo-the-laws-are-written-by-lobbyists/63908 (Accessed: 22 December 2020)

Thrift, N. (2005) *Knowing capitalism*. London: SAGE Publications.

Toff, B. and Nielsen, R. K. (2018) '"I just Google it": folk theories of distributed discovery', *Journal of Communication*, 68(3), pp. 636–657.

Tsesis, A. (2018) 'Marketplace of ideas, privacy, and the digital audience', *Notre Dame Law Review*, 94, p.1585.

Turow, J. (2005) 'Audience construction and culture production: marketing surveillance in the digital age', *The ANNALS of the American Academy of Political and Social Science*, 597(1), pp. 103–121. doi: 10.1177/0002716204270469.

Turow, J. (2011) *The daily you: how the new advertising industry is defining your identity and your worth*. New Haven: Yale University Press.

Turow, J., McGuigan, L. and Maris, E.R. (2015) 'Making data mining a natural part of life: physical retailing, customer surveillance and the 21st century social imaginary', *European Journal of Cultural Studies*, 18(4-5), pp. 464–478.

Tversky, A. and Kahneman, D. (1986) 'Rational choice and the framing of decisions', *Journal of Business*, 59, 4, pp. S251–S278.

Uncles, M. and Lee, D. (2006) 'Brand purchasing by older consumers: an investigation using the Juster scale and the Dirichlet model', *Marketing Letters*, 17(1), pp. 17–29.

Uncles, M., Kennedy, R., Nenycz-Thiel, M., Singh, J. and Kwok, S. (2012) 'In 25 years, across 50 categories, user profiles for directly competing brands seldom differ', *Journal of Advertising Research*, 52(2), pp. 252–261.

Vaidhyanathan, S. (2012) *The Googlization of everything:(and why we should worry)*. Berkeley, California: University of California Press.

Vakratsas, D. and Ambler, T. (1999) 'How advertising works: what do we really know?', *Journal of Marketing*, 63(1), pp. 26–43.

Van Den Driest, F., Sthanunathan, S. and Weed, K. (2016) 'Building an insights engine (cover story)', *Harvard Business Review*, 94(9), pp. 64–14.

Van Dijck, J. (2013) *The culture of connectivity: a critical history of social media*. New York: Oxford University Press.

Van Dijck, J. and Poell, T. (2013) 'Understanding social media logic', *Media and Communication,* 1(1), pp. 2–14.

Varey, C. and Kahneman, D. (1992) 'Experiences extended across time: evaluation of moments and episodes', *Journal of Behavioral Decision Making*, 5(3), pp. 169–185.

Varian, H.R. (2010) 'Computer mediated transactions', *American Economic Review*, 100(2), pp. 1–10.

Varian, H.R. (2014) 'Big data: new tricks for econometrics', *Journal of Economic Perspectives*, 28(2), pp. 3–28.

Varis, P. (2016) 'Digital ethnography' in Georgakopoulou, A. and Spilioti, T. (eds). *The Routledge handbook of language and digital communication.* Abingdon: Routledge, pp. 55–68.

Vega, T. (2012) 'Voters say they are wary of ads made just for them', *The New York Times,* 23 July [online]. Available at: https://www.nytimes.com/2012/07/24/business/media/survey-shows-voters-are-wary-of-tailored-political-ads.html (Accessed: 26 March 2020)

Volkmer, I. (ed.) (2012) *The handbook of global media research*. Oxford: Wiley-Blackwell.

Voorveld, H.A. and Valkenburg, S.M. (2015) 'The fit factor: the role of fit between ads in understanding cross-media synergy', *Journal of Advertising*, 44(3), pp. 185–195.

Vranica, S. (2018a) 'Facebook and Google face emboldened antagonists: big advertisers', *The Wall Street Journal* (English edition), 25 March [online]. Available at: https://www.wsj.com/articles/facebook-and-google-face-emboldened-antagonists-big-advertisers-1521998394 (Accessed: 02 January 2019)

Vranica, S. (2018b) 'Unilever threatens to reduce ad spending on tech platforms that don't combat divisive content', *The Wall Street Journal* (English edition)*,* 11 February [online]. Available at: https://www.wsj.com/articles/unilever-threatens-to-reduce-ad-spending-on-tech-platforms-that-dont-combat-divisive-content-1518398578 (Accessed: 12 January 2019)

Wall, D.S. (2019) 'Surveillant internet technologies and the growth in information capitalism: spams and public trust in the information society', in Haggerty, K. and Ericson, R. (eds.) *The new politics of surveillance and visibility*. Toronto: University of Toronto Press, pp. 340–362. https://doi.org/10.3138/9781442681880-015

Wang, Z. and Tchernev, J.M. (2012) 'The "myth" of media multitasking: reciprocal dynamics of media multitasking, personal needs, and gratifications', *Journal of Communication*, 62(3), pp. 493–513.

Wanta, W. and Hu, Y.W. (1994) 'The effects of credibility, reliance, and exposure on media agenda-setting: a path analysis model', *Journalism Quarterly*, 71(1), pp. 90–98.

WARC (2019) *Adspend database* [Online]. Available at: https://www.warc.com/data/adspend/takeaway-tables (Accessed: 23 April 2019)

Warner, C., Lederer, W. A. and Moroz, B. (2020) *Media selling: digital, television, audio, print and cross-platform.* Oxford: Wiley-Blackwell.

Watkins, E. A. (2019) *Guide to advertising technology* [Online]. Available at: https://academiccommons.columbia.edu/doi/10.7916/d8-9aep-r971 (Accessed: 23 April 2020)

Webster, J., Phalen, P. and Lichty, L. (2013) *Ratings analysis: audience measurement and analytics*. New York: Routledge.

Webster, J.G. (2005) 'Beneath the veneer of fragmentation: television audience polarization in a multichannel world', *Journal of Communication*, 55(2), pp. 366–382.

West, E. (2019) 'Amazon: surveillance as a service'. *Surveillance & Society*, 17(1–2), pp. 27–33.

Westin, A. F. (2000) 'Intrusions', *Public Perspective*, 11(6), pp. 8–11.

Whitaker, R. (2000) *The end of privacy: how total surveillance is becoming a reality*. New York: The New Press.

Whitson, J. R. (2013) 'Gaming the quantified self'. *Surveillance & Society*, 11(1–2), pp. 163–176.

Williams, M. H. (2019) *Data privacy – are you a Google or an Apple?* [Online]. Available at: https://www.warc.com/newsandopinion/opinion/data-privacy--are-you-a-google-or-an-apple/3306 (Accessed: 21 October 2020)

Wind, Y., Sharp, B. and Nelson-Field, K. (2013) 'Empirical generalizations: new laws for digital marketing: how advertising research must change', *Journal of Advertising Research*, 53(2), pp. 175–180.

Wolf, M. J. (1999) *The entertainment economy. The mega-media forces that are re-shaping our lives*. New York: Penguin Random House.

Wolff, M. (2017) *Television is the new television: the unexpected triumph of old media in the digital age*. New York: Penguin Random House.

Wolny, J. and Mueller, C. (2013) 'Analysis of fashion consumers' motives to engage in electronic word-of-mouth communication through social media platforms', *Journal of Marketing Management*, 29(5/6), pp. 562–583.

Wong, C.H., Tan, G.W.H., Tan, B.I. and Ooi, K.B. (2015) 'Mobile advertising: the changing landscape of the advertising industry', *Telematics and Informatics*, 32(4), pp. 720–734.

Wu, T. (2017) *The attention merchants: the epic scramble to get inside our heads*. New York: Vintage Books.

Young, A. (2014) *Brand media strategy: integrated communications planning in the digital era*. New York: Palgrave Macmillan.

Young, M. (2017) *Ogilvy on advertising in the digital age*. London: Goodman Books.

Ytre-Arne, B. and Moe, H. (2020) 'Folk theories of algorithms: Understanding digital irritation', *Media, Culture & Society*. doi:10.1177/0163443720972314.

Zackon, H., Marci, C. and Nuta, N. (2017) *The mind of a reader* [Online]. Available at: http://www.researchexcellence.com/files/pdf/2017-03/id423_cre_the_mind_of_the_viewer_arf_presentation_3.14.17.pdf (Accessed: 24 November 2018)

Zajonc, R.B. (1998) 'Emotions' in Gilbert, D.T., Fiske, S.T. and Lindzey, G. (eds.) *Handbook of social psychology, Vol. 1*, 4th edn. New York: McGraw-Hill, pp. 591–632.

Zuboff, S. (2019) *The age of surveillance capitalism: the fight for a human future at the new frontier of power.* New York: Public Affairs.