University of Kent

MSc Computer Science

A dissertation submitted in fulfilment of the requirements for the degree of MSc Computer Science in the School of Computing, Faculty of Sciences

---

A Critical Analysis of Facebook-Cambridge Analytica scandal, Regulation of Data protection legislation (UK's DPA 2018 & GDPR) and the future of Data-Driven Political campaigns.

Author:
Neelansh Balhara

Supervisor:
Dr. Carlos P. Delgado

October 8, 2020

Table of Contents

1 Introduction

This study is built around one of most significant presidential elections in the past decade the 2016 US presidential elections; addressing the reasons for why the atmosphere surrounding the elections was cynical and tense. While following through the elections during 2016, one of the most important investigative event took place shortly after the election results were announced. The UK's information commission set out an investigation into the matters of data analytics and political campaigning. The ICO had little clue about what this investigation will turn into not long after they initiated it. What turned out to be the biggest data protection investigation ever conducted the information commission focused their resources towards two corporations, Facebook.Inc and Cambridge Analytica. The result, numerous accounts of data breaches which concerned the personal information of over 87 million users worldwide, out of which 1 million users in the UK were affected as their personal data was used for political purposes without any notification of intent to do so by the social media giant. 37 million American's data was unlawfully taken by the data analytics firm Cambridge Analytica which participated as the political consultants to Donald Trump's presidential campaign. The grounds for investigation were strong and it did not take a long time before this scandal became an international news.

In the context of this thesis, my primary aim is to understand how the involvement of data analytics into political campaigning can undermine democratic practices. The sections that follows, paints a very broad picture of the scale at which data analytics is involved with political parties as i take one of the most lucrative political consultancies SCL Elections limited and highlight their political research operations since 2003 in India, which is one of the many countries SCLE[1] has its claws into. By doing so, I intend to give the scale of which data analytics is involved with politics outside the EU and USA for the reader to take into account that this issue is not subjected to just one nation, it is by all means a global issue. Furthermore, two of the building blocks for this study is presented in the form of an analysis of the role that political parities play in the digital campaigning ecosystem as they sit at the top of the pyramid as clients for data analytic services. The importance of this section is to highlight the responsibilities political parties must embody as data controllers to serve by example, the importance of safeguarding their citizens personal data. In this analysis, I expand on the involvement of Cambridge Analytica and the role they played in Trump's campaign which involved designing political adverts based on the processed personal data they acquired unlawfully from Facebook's platform. In addition to their involvement as political consultants, the general tone of the elections boiled down to character attacks by both the sides which were regularly mediated throughout the multiple domains of media.

The second half of this study majorly deals with data protection legislation in the UK and EU's GDPR, by highlighting and providing a comparative analysis between the data protection principles in the DPA-1998 and GDPR, I aim to familiarise the reader with their rights as data subjects, along with the standards by which data controllers and processors must handle your personal data. Both the parties involved in the Facebook-Cambridge Analytica scandal in my estimation, had no regards for the legislation they are meant to abide by. This section dives into the notion of valid consent, political interest in personal data and legal actions taken by the ICO towards Facebook, Cambridge Analytica and SCL Election Ltd.
This scope of this breach and reactionary responses from Facebook regarding the ICO's investigation were rather vague and lacked the seriousness that the this issue demands and expects from the tech giant. In the section that follows, an analysis is presented of how the political campaigns in near future may look like based on the context of this study and the proliferation of data. The main challenges for data protection authorities are yet to come as the data protection will face major obstacles in terms of opacity of algorithmic processes, transparency and big data analytics, the benefits of these advancements should not be achieved at the expense of data privacy rights.

[1]SCLE Ltd. is also the parent company of Cambridge Analytica

2 Network Democracy: Motivation and Research Question

> "One of our deepest liberal democratic institutions is that generalized advance in our ability to gather and share information, and to communicate with one another, invigorates democratic politics."
> -Darin Barney (The Network Society)

Democracy, in whatever form it exists is a deeply communicative brand of politics as it demands multidimensional exchange of information and views. It requires dialogue and a public sphere to breathe, in which citizens can participate in practices that define them as citizens and their societies as democratic that involves dissemination of information; expression and consideration of opposing viewpoints; critical debate on issues of common public concern; scrutiny of public authorities and policies. However, as historically observed the relationship between mass communication technologies and democracy has never been simple and neither it is the case here. The most important example in this context is outlined by Bruce Bimber in an article documenting the absence of statistical evidence linking internet usage independently to increased political engagement.

> Opportunities to become better informed have apparently expanded historically, as the informational context of politics has grown richer and become better endowed with media and ready access to political information. Yet non of the major developments in communication in the 20th century produced any aggregate again in citizen participation. Neither telephones, radio, TV not the internet exerted a net positive effect on participation, despite the fact that they apparently reduced information costs and improved citizens' access to information. [2]

The association of data analytics providers and political parties seems a natural outcome of exponential development in information technologies and computer sciences in terms of large scale implementation, reliability and effectiveness of both hardware and software. It also proves to serve a useful function in political campaigning as it provides highly sophisticated services for campaigns to get invaluable insights to their voter population which accounts for its wide range deployment by data driven political campaigns.
This almost integrated association between political campaigns and data analytics have never been observed before have only emerged out of which seems to be, two feasible reasons i.e. primarily, rapid growth of information and technologies associated with information (whether information generation, storage, maintenance, development and/or safeguarding.)
Secondly, the applicability, adaptability and accuracy data analysis techniques in political campaigning, which have been strongly displayed during the 2016 US presidential elections.
The nature of that data which is exploited by certain data analytics firms is mainly personal data, involving various personality aspects of the user. In order to gain personal insights about the voter population and making room for the propagation false interpretations of an ideology, misinformation. For example, in an capitalist economy, to be pro job seems an obvious inclination towards success or meaning some would say, yet stating that alone is not enough to differentiate one political candidate competing with another since the above mentioned claim is accepted widely and without hesitation. However, what would give one candidate preference over another would be their ability to analyze their target voter population, profile them in an order of personalities profiles. The classification criterion itself would be based on personality traits and is synthesised from personal data collected through various different sources.
Given the argument about the inevitability of rapid information growth and the association of data analysis and political campaigns and how this association changed, entirely the type of politics being conducted, almost every political campaigns now constitutes and heavily rely on a dedicated IT cell with a primary function of advertising about their political manifesto to very specifically identified groups of people.
It is not to say that the sophistication of these techniques have always been employed to undermine a

[2]Bimber, Bruce. "The Internet and Political Transformation: Populism, Community, and Accelerated Pluralism." Polity, vol. 31, no. 1, 1998, pp. 133–160. JSTOR, Available at: [www.jstor.org/stable/3235370.]

democratic institution. The transformative power of big data analytics have been observed in academic research, health-care, social sciences, cosmology and mainstream economics to name a few. One of the primary aim of this analysis is to investigate the impact of data analytics and big data in regards to political campaigning. One of the primary driving factors to start this assessment was the association of Cambridge Analytica with Donald Trump's 2016 US presidential elections and the nation wide political research in 10 major Indian cities by Cambridge Analytica's parent company SCLE Ltd. These firms provide a wide range of political consultancy and services to the highest bidder, this shift of power allows data analytics to penetrate deep within the sanctity of fair elections in a democratic society; as it is now convenient for them to empower certain campaigns and put their key players wherever they want.
On an intuitive level this tight association raises some concerns, for instance, the polarization we observe throughout major developed countries, at least on a political level, Trump's victory in 2016 US presidential elections or the triumph of BJP[3] in 2014 Indian national elections, with an astounding majority of seats 282 out of 543, 166 seats more than in the previous elections. Cambridge Analytica was the political consultants for these campaigns. One of main question this thesis outlines by an extensive analysis of the Facebook-Cambridge Analytica scandal as a central pillar is, do these campaign strategists be able to fully utilize analytics of big data to misinform voters via pushing the public discourse in illegitimate ways? Will the politics of persuasion boil down to character attacks, spread racist, sexist and ill-informed ideas echoing in very specific echo-chambers as this discourse further polarize nations?
This argument is not in any way rooted in a reactionary suspicion towards big data analytics, but rather on an academic curiosity to examine the course of actions that took place within the past decade that strengthened the association of data analytics in political campaigning.
Data Analytics and the revelation of big data have undoubtedly, been proven highly beneficial to democracy just as its resourcefulness is widely accepted and acknowledged in the scientific community. Yet there is little evidence in favor of how these techniques have fed into the democratization of societies or to what extent they are used by governments to empower the democratization of societies; again it is somewhat morally neutral. What is of up most importance is the strengthening of regulatory authorities along with the voluntary compliance of data controllers and processors in order to continuously examine this association of politics and data analytics to make sure that it develops towards a fruitful end for everyone involved.
Going back to addressing the scope of this study, it will primarily revolve around the association of data analytics and political campaigning. Democratic elections constitute a fundamental (however not exclusive) aspect of democracy. Data analytics, due to its effectiveness in generating useful controls and services constitute a great deal of utility value, specifically in political campaigning. The research question can then be outlined in the following manner,

> Does the rigorous involvement of data analytics in political campaigning undermines transparency in democratic elections?

Secondly, since the enforcement of EU's General Data Protection Regulations since 31 May 2018 in the UK the second half of this analysis would be based on fundamentally the analysis of this legislation. By this analysis, I aim to underline the importance of strong data protection legislation as it seems to be the only tool at hand while dealing with massive corporate entities; along with the transnational nature of data that serves as the key element to the digital campaigning ecosystem.

3 Methodological Approach

The nature of the questions raised in the section above requires a philosophical assessment initially. The key challenge during compilation of this thesis was following up and assessing both sides of the story. The investigation into the whole Facebook-Cambridge Analytica scandal that spanned over 3 years and more, since the December of 2017 till the issue of monetary penalty issued by the ICO to the tech giant during October 2018 was the biggest data protection investigation ever conducted. What was clear since

[3]Wikipedia contributors. (2019, August 16). 16th Lok Sabha. In Wikipedia, The Free Encyclopedia. Retrieved 21:21, August 20, 2019, Avaulable at:[https://en.wikipedia.org/w/index.php?title=16th_Lok_Sabha&oldid=911131005]

the early developments into the investigation was the involvement of a web of corporations in and out of ICO's jurisdiction. This whole investigation set in motion a serious public concern into the matters of online privacy, data protection and the limitations in the current data protection legislation. In the scope of this analysis which, fundamentally is an assessment of one of the most significant data breaches of the 20th Century. However it is the series of questions that surrounds the breach that makes it so significant, this analysis will address till some extent uncover broadly and extensively the changed nature of political campaigning with the association of data analytics and what issues those changes lay before us.

This analysis requires three things which are fundamental to it. Firstly, the analysis of Democracy in a networked world on at least a conceptual scale, this conceptual analysis serves as the foundation to this thesis. A second assessment will be based around the role of data analytics in the digital campaigning ecosystem, this will be primarily a factual analysis of how campaigns are tilting towards the lucrative model of data analytics; provided a detailed inspection of the investigation conducted by UKs Information commission into the Facebook-CA scandal.

Following that an assessment on the current legal protection in place to deal with any misconduct relating to personal data which is the EU's GDPR. The motivation of providing this analysis is to in some sense it to familiarize the reader with their rights as data subjects along with how entities operating with citizens personal data should operate in the domain of protecting those rights.

Thirdly the discussion on democracy, data analytics in political campaigning and data protection legislation require to be combined into a coherent theoretical assessment into what the future of political campaigning holds. The methodological approach is by its nature twofold, as the primary assessment in deductive where I present a research question based on the theoretical discussion preceding it, following it, empirically studying relevant, suitable cases. This pattern of assessment is followed throughout this thesis. My aim regarding this study is to present a case of how prevalent the theoretically deduced primary research question is in real life.

In this context, this allows me assess thoroughly my research question though an inductive analysis, at the same time providing me with a window to reformulate my initial conclusions. However in the domain of this thesis my primary focus would be on the former. I invite other researchers to engage in the analysis of the latter. In summation, this analysis should be viewed as an assessment of a series of really significant events that affected the whole digital campaigning ecosystem, online privacy and data protection legislation in the UK. This analysis will in my estimation will allow me to investigate, to some extent what does the future of political campaigning looks like.

3.1 Online Political Campaigning.

Throughout the history of electronic mass media's availability in extensive public domain, there exists an undeniable political interest in its captive power. The brief history of the early internet can be distinctively divided into two eras, "Web 1.0 and 2.0" (Vergeer, M 2013) . Although, e-campaigning was a common thing during the Web 1.0 era but it wasn't until the development towards Web 2.0 took an exponential rate of growth it completely changed a vast majority of entities associated with it, as a result e-campaigning witnessed a tremendous growth. The facilities provided by the framework of Web 2.0, they lie in the network's capabilities to grow and include vast majority of population. What is noticeable and particularly relevant to this thesis is that the political campaigns which significantly facilitate and in some cases exploit the ability to reach people via social media. The social media outlet allows politicians to develop highly personalized and individualized political campaigns and to what extent that feeds into the democratization of societies and nations is up for debate. Past studies regarding the problem voter participation suggests that, "social utility is particularly important and a strong motivation for information seeking,"[4]

Social utility serves as a leading cause for citizen to connect with politicians followed by entertainment. This raises a serious series of questions regarding the impact of online campaigning on political participation of the citizens, does e-campaigning have facilitated a means of countering the steady decline in

[4]Vergeer,M. (2013). "Politics,elections and online campaigning:Past,presentand a peek into the future." New Media Society, 15(1), 9–17. Available at:[https://doi.org/10.1177/1461444812457327]

citizen participation in politics? Or is it the case that during the course of the transformation from Web 1.0 to 2.0 and the significant presence of social media activities, the fundamental goal of engaging public's participation into politics is lost? These are few of questions whose resolution require an extensive analysis. However, the popularity of social networking sites and platforms cannot go unnoticed.
A significant part of the political communication nowadays takes place over a variety of outlets, weather as a conventional multi-cast political debate on the TV, over social media outlets since political campaigns and ideologies now have a medium to represent themselves as personalized social media accounts via which they can interact with each other and as well as with the vast majority of public who also share the same cyberspace as them. With the emergence of Big data and easier than ever aggregation of various datasets for very sophisticated analysis, the very complexity of this discourse of political communication also gives birth to various extremely sophisticated and ways to study political behavior and online public opinion. Moreover, the advances made in the ways of collecting, analyzing and mining useful knowledge from online data have been exponential, from various APIs provided by online platforms, scraping softwares like Thelwall's Soscibot, Ackland's uberlink or Connected Action's NodeXL to Tracking cookies, growth of IoT and highly accurate targeting methods have made politics and political engagement interesting not only for citizen but also for political parties and campaigns, data controllers and broker, social media giants (such as Facebook, which incorporates powers of data controllers too) and specially for scientists and psychometric analysts. In this context, the emergence of a digital campaigning ecosystem seems natural in which these entities interact and generate both useful knowledge and capital of national as well as international interests.
The ability to communicate and engage with voters is an essential democratic practice. It is an important part of any political campaign's success. It is in the benefit of political parties to maintain an open voter engagement since without it, parties would not be able to priorities citizens' concern and neither would their polices in order to deal with those concerns. One of the results observed as a direct result of improper voter's engagement and participation is the decline in influence citizens have over policies and democratic changes creating an atmosphere of tension among governments and citizens. Throughout the history of political campaigning political parties have incorporated a number of communication methods to engage with their public, that being said it is also true that these methods and processes have developed over decades, considering since the development of mass communication and broadcast technology in the 1930s to the age of Television in the 50s enabling political parties to speak directly to large group of voters to the early internet political discussion forums in the 90s; these advances developed over time in line with the technological advancements.

> "The post-war period also saw the growth in advertising and market research techniques, which enabled political parties to better understand voters' concerns, and to shape messages accordingly through political advertising – with the use of large advertising agencies becoming widespread during the 1980s and 1990s. The advent of telephone and email canvassing by political parties and campaigns enabled direct individual contact with voters on a mass scale."[5] (ICO, July 2018)

Now, in the age of social media and Big data as they completely changed the dynamics of political campaigning, enabling high end data analytic firms such as SCLE Ltd, Cambridge Analytica and Aggregate IQ to empower political campaigns, globally with rich, accurate demographics of their voters. These no shortage of political interest in these services and their demands have reach a new high. Political parties in the EU and globally spend a significant amount of financial resources. In an 2017 analysis on the expenditure by political campaigns in the UK on Facebook advertisements done by the BBC Political Research Unit, 'In an over all break down, parties spent almost £40m durnig 2017 general elections, with a leading expenditure by the Conservative and Unionist Party: £18,565,102 closly followed by the Labour Party: £11,003,980.'[6] In addition to this, the parties spent around £3m directly on

[5]ICO, Democracy Disrupted?, 11 July 2018, Section 1.1 Avaiable at:[https://ico.org.uk/media/2259369/democracy-disrupted-110718.pdf]

[6]Joey D'Urso, BBC Political Research Unit, Who spent what on Facebook during 2017 election campaign?, 31 March 2018, Available at: [https://www.bbc.co.uk/news/uk-politics-43487301]

Facebook and that capital was not evenly distributed. Figure-1[7] below shows that the amount of money invested in Facebook advertisement by the Conservative party was twice as much as all other parties combined. Political interest in taking advantage of new techniques in the frontier of digital campaign-

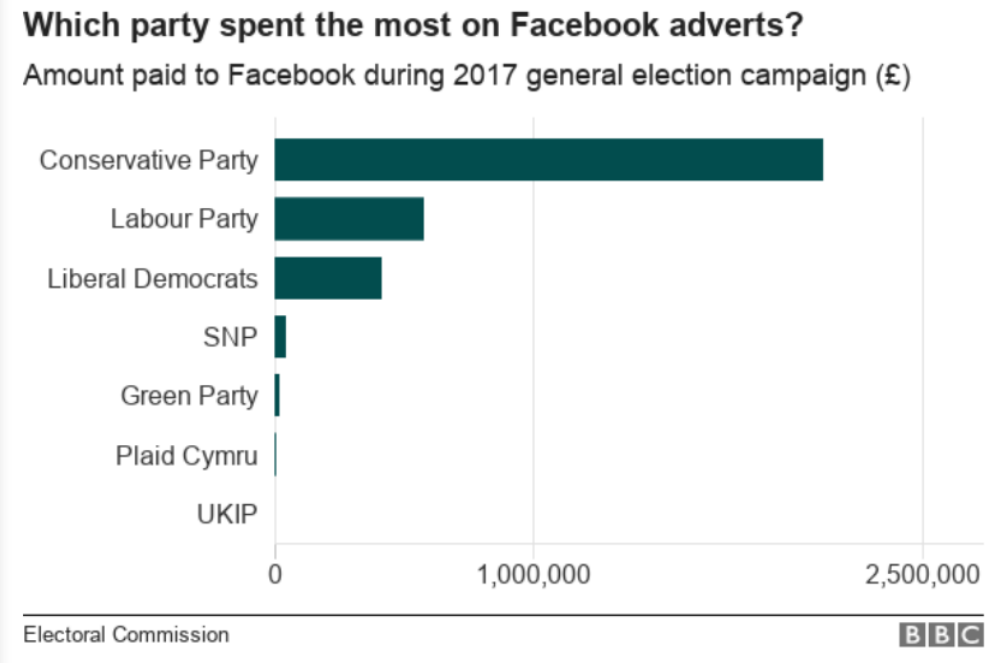

Figure 1: Amount of money paid to Facebook during 2017 general election for running political adverts by seven election campaigns.

ing is understandable. Elections all over the world, specially in developed and developing nations are getting increasingly datafied. As of its natural byproduct there is an emergence of highly sophisticated and accurate advertising and marketing techniques provided primarily by private data contractors and analytic firms, their services include high-end audience segmentation and targeting to political parties all over the world. For instance during India's 2014 elections, Cambridge Analytica's parent firm proposed a massive political machine for India's major political parties. SCL Elections Ltd proposed a research based operational centre as they have claimed to have set-up in Indonesia, Kenya, Thailand and the United Kingdom. As mentioned in the report the OpCentre according to their planning would incorporate a wide range of "support workers, senior lawyers and media-monitoring professionals" providing more technical guidance to party workers.

> "According to documents obtained by Quartz, SCL claimed to employ as many as 300 permanent staff and 1,400 consultant staff prior to the country's 2014 general elections, which were swept by the Bharatiya Janata Party (BJP) of prime minister Narendra Modi. Ahead of the elections, SCL targeted India's major parties by pitching a setup of "the most sophisticated political research and data hub" in the country."[8]

3.2 SCL Elections Ltd Indian operations.

SCL is the parent of a group of companies. Who exactly owns SCL and its diverse branches is unclear, thanks to a convoluted corporate structure, the type seen in the UK Companies House, the Panama Papers, and the Delaware company registry. Some of the SCL offshoots have been involved in elections from Ukraine to Nigeria, helped the Nepalese monarch against the rebels, whereas others have developed methods to influence Eastern European and Afghan citizens for NATO. And, in 2013, SCL spun off a new company to participate in US elections: Cambridge Analytica. (Green and Issenberg, 2016) In the scope

[7]BBC Political Research Unit, [https://www.bbc.co.uk/news/uk-politics-43487301]

[8]Itika Punit, Cambridge Alaytica's parent firm proposed a massive political machine for India's 2014 election, 29 March 2018, Quartz Infrastructure and capabilities, Available at:[https://qz.com/1239561/cambridge-analyticas-parent-firm-proposed-a-massive-political-machine-for-indias-2014-elections/]

of this section, I present the hold of SCL Elections Ltd thoughout India. SCL Elections Ltd have been operating in 10 major Indian cities since 2003.[9]In the obtained documentation by Quartz, they paint a much bigger picture than that of just a political operation centre. During a series of event preceding the

When	Where	What
2012	Uttar Pradesh (UP)	A caste census in on behalf of a national party
2011	UP	Statewide (200 million people) research campaign to identify voter caste by household.
2010	Bihar elections	Electoral research and strategy for the Janata Dal (United)
2009	National elections	Managed campaigns of a number of Lok Sabha candidates
2007	UP	Full political survey on behalf of a major party
2007	Kerala, West Bengal, Assam, Bihar, Jharkhand and UP	Research communication campaign to counter the recruitment into, and support for, "violent Jihadism" in six states
2003	Madhya Pradesh elections	Psephological study and opinion polling for a national party to identify swing voters
2003	Rajasthan elections	Assessed a major state party's organisational strength, and nature of the voting population and the attitudes and behaviours of politically active individuals within the state

Figure 2: SCL Elections Ltd. Indian operations since 2003 to 2012

fallout of Cambridge Analytica, the former research director at Cambridge Analytica and the whistle blower of the whole Scandal Christopher Whylie revealed via a series of leaked documentation regarding both Cambridge Analytica and SCLE Ltd.[10] On March 28, Whylie posted screenshots from the same documentation obtained by Quartz highlighting these same projects. During almost a decade, SCLE Ltd conducted a number of surveys through the Indian territory. What is unanimously agreed upon is the fact that unlike relatively traditional methods of political campaigning, these techniques are by their design and nature more opaque. Specially in more volatile nations, specially in the worst largest most religiously divided democracy, the long term repercussions of employing audience segmentation techniques can be unpredictable. Highly targeted messages are often captured in echo chambers online, where the audience may not even get a chance to hear the other side of the issue which puts a limit on the perspectives that issue or argument could be looked at. Voters may at times not even understand why they are receiving particular messages or the provenance of the message. Free, fair and transparent elections are the bedrock of democracy and are highlighted in the human rights law, the EU's Data Protection Supervisor's report on online manipulation outlines the negative consequences of not conducting electoral practices up to the standards of the principle of electoral transparency.

[9]10 Indian cities of SCLE Ltd. offices: Ahmedabad, Bengaluru, Cuttack, Ghaziabad, Guwahati, Hyderabad, Indore, Kolkata, Patna, and Pune.

[10]Twitter: @Chrisinsilico, Available at: [`https://twitter.com/chrisinsilico/status/978921850448371715`]

"The principle of electoral transparency is not met if the voters have no freedom to seek, receive and impart information about the process and the candidates, including about the source and spending of financial support received by a candidate or a party. These rights are also therefore challenged by online manipulation"[11]

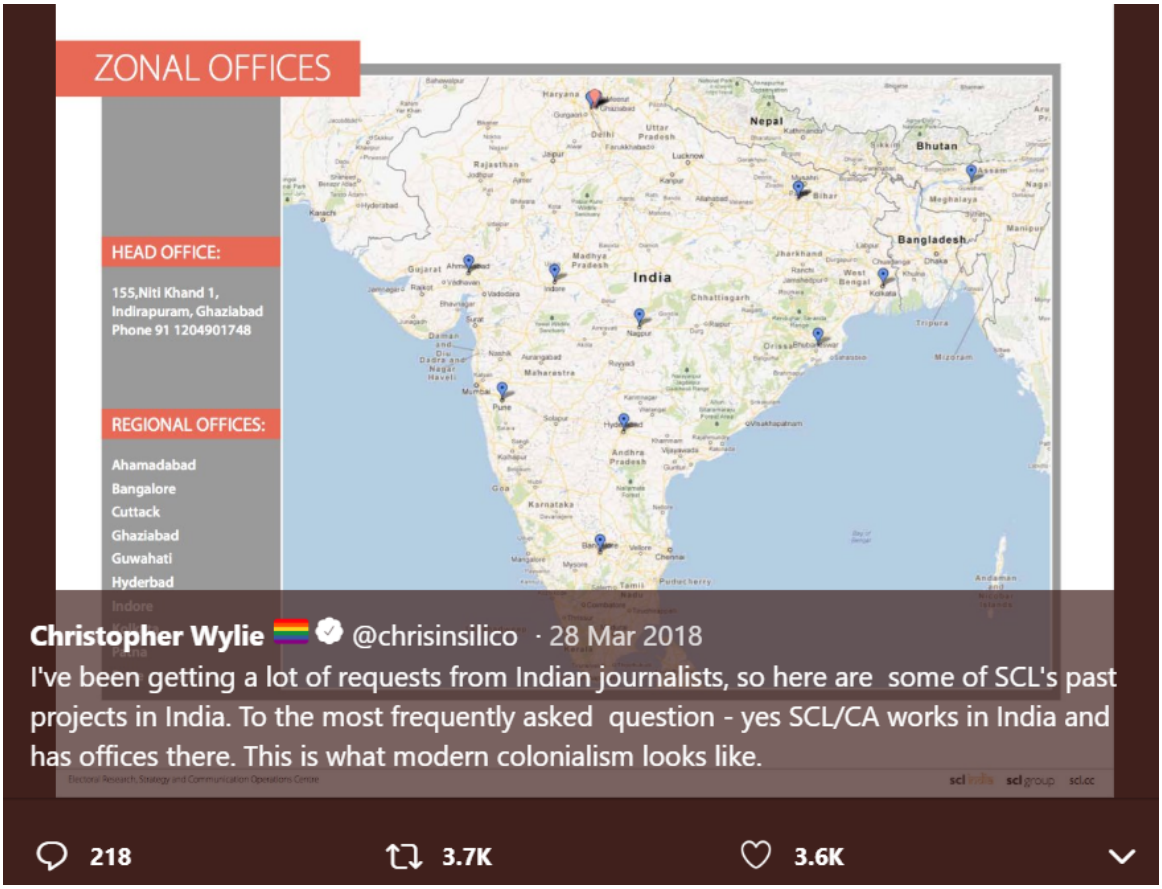

Figure 3: A post on twitter by former research director Christopher Whylie outlining SCL Elections Ltd.'s past projects and operations in India from 2003 till 2012.

The UK's information commission emphasizes on the importance of operating on a level playing field when accessing the electorate and strongly suggest political parties and campaigns to provide the voters with the full spectrum of their ideology and information regarding political messaging and who the are faces behind their messages. As the rules completely change when the playing field is cyber space and data is the crucial element for these entities to thrive and operate. The game however is still the same. In the same context the menu of political services SCLE Ltd provided to the Indian parties for the 2014 election under the name of "examples of past research", the firm conducted several wide range of surveys which raises question regarding domestic and international security as a foreign political consultancy firm with a global clientele conduct such massive surveys collecting highly personal data of over millions of Indian citizens. In the documentation acquired by the Quartz SCLE Ltd undertook several researches.

These services included a detailed analysis to quantify independent (for whom people intend to vote) and dependent (why people intend to vote/not vote for a particular candidate) variables under the name of Behavioural polling. This information serves as fundamental building block to build a voter profile which can be employed by campaigning team for maximum usage, weather that is highly targeted political messages or advertisements. The documentation cites the examples of voter migration in 2007 Lucknow East elections where according to their research,

> BJP (Bharatiya Janta Party) has the most loyal voters, followed by the INC (Indian National Congress) with 70% of (the) 2007 voters remaining loyal. The BSP (Bahujan Samaj Party) and SP (Samajwadi Party) both lose voters to the BJP and INC (Indian National Congress), respectively.

[11] EDPS Opinion on online manipulation and personal data, March 2018, Available at:[`https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf`]

Another research conducted by the firm took place in two areas of the largest Indian state Uttar Pradesh, in Kairana and Kanth respectively. This aim of the research being the influence of Caste in electoral participation of the citizens in a region and then shaping messages according to the obtained results. Their Candidate research included the objective to identify apathetic voters, floating voters, desire for change within a population, the best medium for transmitting campaign messages and weather **a positive campaign will be most effective or a negative one.** These operational researches were being conducted in highly populated areas of major Indian states long before the arrival of General Data Protection Regulations in the UK. There is no denying that digitization of society and the economic structure of many nations is having somewhat of a mixed impact on the civic engagement in democratic processes and activities. Various western data protection authorities including the EDPS[12] have for several years argued for greater collaborative efforts between data protection authorities and other regulators globally; along with building up to a collaborative framework where DPAs have an open dialogue on enforcing data protection regulations on political parties, campaigns and all multinational organizations that deal with personal data. It has been stressed upon via various regulatory authorities that data protection is not just a matter of breaches and quantifying loss, it is a matter of keeping a democratic society healthy by safeguarding citizens' trust in democratic elections. Elizabeth Denham, UK's information commissioner under a committee hearing held by House of Commons stated, when asked about the Canadian connection to the Facebook-Cambridge Analytica scandal specifically raising concern regarding the revelation made by Christopher Wylie who stated that SCL Election Ltd found it convenient to set up a number of player in different jurisdictions. It was easier for a number of reasons for them to run these campaigns. Upon the DCMS investigation that highlighted the jurisdiction issue and what can be done in order to address as despite the best efforts of the Parliament.

> "I think that with digital campaigning, big-data politics, big-data commerce and the fact that personal data is the asset that is driving the internet, it is extremely important that data protection regulators have the ability to reach beyond their borders—extraterritorial reach. In the UK and the EU, we are now able to follow the data in that way. But it is also very, very important that we have strong enforcement powers to compel documentation and to reach into the cloud, in a digital search. That is very important."[13]

It is crucial now more than ever that international regulatory co-operation networks and relations are strengthened. In the long term, regulatory authorities might have to lean towards multilateral treaties. Due to the transnational nature of data as it have no borders now, international co-operation need to come together towards a global solution. Moreover, the requirement of strong enforcement tools to move online platforms in the right, lawful direction so that they take the people with them in the services they offer; as Facebook remains concrete to their business model seriously stating no changes will be made to their business model and their practice unless there is a legal order to do so.

3.3 Network Politics

This section emphasizes on the association of network technologies and politics and further expands on the transformation of politics due to the rapid technological developments that seems to transform everything it touches. In the early 1990s up till 2000 there were a lot of speculations and prediction regarding the relatively rapid growth of the early internet which was hyperlink based, majority of the information on the early internet was in hyperlink form, along with an idea of a free cyberspace for peer reviewed information sharing and communication.

> because current political systems are still based in organizational forms and political strategies of the industrial era, they have become politically obsolete, and their autonomy is being denied by the flow of information on which they depend. (Castells 1997: 312).

[12]European Data Protection Supervior

[13]Oral evidence: Disinformation and 'fake news', HC 363, 27 Nov 2018, House of Commons (Digital, Culture, Media and Sport International Grand Committee), Q4284, Available at:[http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/digital-culture-media-and-sport-committee/disinformation-and-fake-news/oral/92924.html]

This implies that the new politics is thereby a politics of struggling over information management and control space constructed by prevailing media of communication as a necessary precondition of access to more material form of power.

> Politics, in this view, is a contest to define parameters of public discourse, and the symbolic and cultural codes through which norms and expectations are expressed and circulated. (Barney 2010).

A minimum condition of political action, intuitively, is access to, and presence and/or representation within, the arenas (i.e. mass communication media) in which these battles are engaged, It is also for the same reason that those who are systematically denied access to advanced information and communication media, or whose access to them is limited to passive consumption of commodified content, are not only economically disadvantaged in the network society, but also politically disenfranchised.

This *digital-divide* is, consequently, at once a technological, economic and political divide, a divide which sets the terms of access to citizenship itself, both within technologically developed regions and between wealthy and impoverished areas in the global system.

In this context politics simply boils down to a profession of public relations instead of a practice of public judgment and action. This information politics manifests itself in terms of trading in the complex crafting and circulation of highly coded messages, datasets etc by employing sophisticated technologies of information gathering and dissemination. (Barney 2010) Observed historically, is an uneven distribution of access to these resources. Politics of persuasion and transparency have always been two fundamental agents for stability in a democracy and therefore democracy is centred upon them.

> Today, the ability to shape and move public discourse through communication remains crucial to the exercise of power and influence. Barney(2010)

This manipulation of public discourse is rooted deep in the ancient origins of this form of government. Today, the ability to shape public discourse through communication remains crucial to the exercise of power and influence. Communication, based on advance network technologies transformed the way of conducting politics, however the fundamental action of shifting public discourse remains intact. In relation to the exponential growth of network technologies (as we shall see throughout this thesis), the term 'new politics' seems justifiable, in doing so we can expand on the role that new information and communication technologies play in relation to new politics.

3.4 Informational politics of the network society.

If we follow the character of the 'new politics', and pay attention to the role that information and communication technologies play in relation to it, a strong case can be made or profitably discussed in relation to the politics of globalization. According to the model of informational politics, the political contest over globalization is best interpreted as an instance of politics of signification, a struggle over the cultural codes through which the sign "globalization" enters and circulates in public discourse. Barney presents two fundamentally important questions in this regards,

> **Does/will globalization mean extension of the rights and freedoms associated with liberal democracy, a sharing of the prosperity associated with market capitalism, rising standard of living, enhanced intercultural understanding and harmony, international peace, solidarity and global democracy? Or does/will it mean the end of national self-determination and autonomy, the triumph of unaccountable transnational corporations and institutions over democratic governments, global dominance of American cultural commodities, a deepening of the dependency and exploitation of the developing world, environmental degradation and an assault on working people?** Barney (2010)

Clearly, the first "set of meanings is the one preferred by the forces of transnational capital that stand to gain most from globalization on the neo-liberal model; the latter set if meanings is preferred by the multifaceted transnational social movements that has risen in opposition to these forces." (Barney 2010)

It is crucial in this sense to think about what the resultant outcome of these colliding conflicts depends upon, i.e. which will ultimately settle in as definitive in popular discourse and the public imagination depends, in the paradigm being discussed throughout this thesis. According to Barney, on the outcome of an ongoing struggle over cultural definition that is being waged in the global circuits of mediated communication. **"Who will succeed in branding globalization?"** This is the informational politics of the network society.
Now further highlighting the crucial role of network technologies in this context, it is important to highlight that these technologies provide resources to both sides, as they provide the infrastructure for the various economic flows of transnational capitalism in the form of better and faster communication facilities which support multiple dimensions/formats of information and media. Their deterritrorializing effects undermine the ability of national governments that are still fundamentally traditional in this sense, democratically accountable to the public interest, to impose limits on these powerful economic actors.

> It is of no surprise that these same network technologies have also provided an opportunity and means for transnational elite to consolidate their control over the global mediascape, through a dramatic concentration of ownership , through horizontal and vertical integration, across media-platforms, and across the content/carriage divide. (Compaine and Gomery 2000; McChesney 1999; D. Schiller 1999.)

This today can be observed by analyzing both the expansion and the market share of traffic on the web. Over the past six years since 2013, Google and Facebook now exercise direct influence over 70%+ of the internet traffic[14], their dominance over the web is undeniable. Mobile internet traffic (due to the rapid development of IoT technology) is now the majority of the traffic worldwide, "alone in Latin America, Google and Facebook services have had 60% of the mobile traffic in 2015, growing to 70% by the end of 2016".[15] The primary source of access to their services or networks is mobile devices. Another case can be made by emphasizing on the growth rate of the traffic for both Facebook and Google since 2014 among the media websites. Prior to 2014, Search Engine Optimization (SEO) was a common practice among web Developers to improve their site for Google searches, since it accounted for approximately 35% of traffic, while 50% of the traffic came from various other places on the web. "Over the years 2015-2017, traffic from Facebook grew approximately 45% surpassing the status that search traffic had". The media depends on both Google and Facebook for page views since it's the majority of their traffic. The Figure 4, shows the combined percentage of traffic referred by these social media tech giants combined.[16]. Moreover, data indicates that Facebook has dramatically improved its dominance on the Web while at the same time Google Search hasn't changed significantly. The factors contributing to Facebook's success in this race are quick interesting. Before 2014, both the giants had a portfolio of multiple web services.Google tried breaking into the social market, "first with the Google Wave, then Google Buzz, Orkut and Google+. In total, Google has acquired 18 companies from the social media category, while Facebook was competing in the search market via Bing, in partnership with MSFT."[17] One of the most crucial things that happened during the year 2014 was Facebook's self recognition of itself to focus primarily on social market. During February of 2014 Facebook bought WhatsApp for 11 times the price Google bought Youtube for, roughly equating to 19 billion U.S. dollars.

[14]J. Clement, *Mobile internet usage worldwide.* Statista, Available at:[https://www.statista.com/topics/779/mobile-internet/]

[15]André Staltz, *The web began dying in 2014, Statistical review*, Available at:[https://staltz.com/the-web-began-dying-in-2014-heres-how.html]

[16]Allie VanNest, Facebook continues to beat Google in Sending Traffic to Top publishers. December 4, 2015, Parse.ly, Available at: [https://blog.parse.ly/post/2855/facebook-continues-to-beat-google-in-sending-traffic-to-top-publishers/]

[17]André Staltz, *The web began dying in 2014, Statistical review*, Available at:[https://staltz.com/the-web-began-dying-in-2014-heres-how.html]

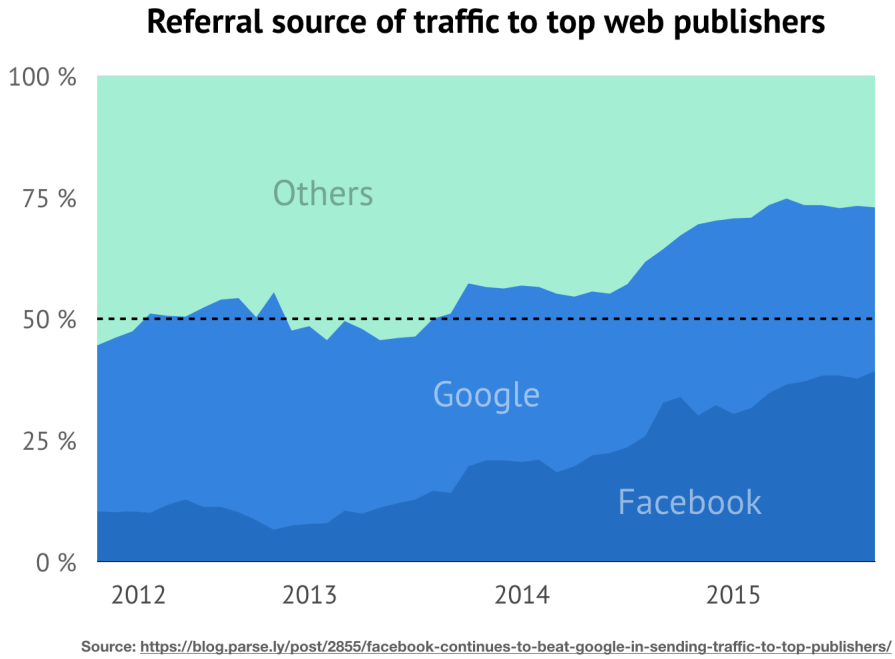

Figure 4: Parse.ly's latest Authority Report revealed that social media not search, brought more readers to the top news stories of 2015. The quarterly review of the top sources of referral traffic came from social media sources.

Following that year in December, Facebook canceled its Bing partnership with Microsoft. As a result, the user retention on Facebook.com grew with a steady pace, through its four simple products, Facebook, WhatsApp, Messenger and Instagram. Facebook had become the social superpower. "With almost 2.4 billion monthly active users, Facebook is the most popular social network worldwide."[18] Facebook alone accounts for over 42% of monthly social media visits and this appeal is not just based on its social platform but also and very much on its strong mobile integration and mobile messaging capabilities. In 2018, the social giants annual revenue amounted close to 55.84 billion U.S. dollars, the majority portion of the income stream generated via advertising. The general prediction of the number of daily users of social networks by 2020 is approximated to 2.95 billion people, majority of this growth is assumed to come from mobile devices, (evidence of which is overwhelming, According to April 2019 data, the global mobile population amounted to 4 billion unique users[19]) as emerging markets catch up on online connectivity.

In order to capitalize on this, Facebook also released a data-efficient mobile app known as Facebook Lite, which is a light-weight version of its conventional mobile app is all set to function on slow internet connections; primarily designed to be deployed in developing nations such as India and the Philippines, ensuring a most definite position for the social network in the local online culture. Likewise, Google in early 2014 started reorganizing itself to focus on Artificial intelligence only. During Janurary of 2014 Google bought DeepMind for 500 million U.S. dollars following the shutdown of its social media platform Orkut. The shift towards Artificial intelligence by Google was in some sense transformative, Google stopped seeing future in the simple search market and announced its migration "from Search to Suggest." It is highly speculated that even though Google is currently slightly behind Facebook in terms of its growing web dominance. Due to their technical expertise, vast budget,influence (both social and political) and vision, in the long haul its AI assets will play a significant role on the internet. Observed today, is a saturation of these tech giants with their orthogonal dominance of parts of the web, this leads to less diverse competition in the internet market space.

Looking back to the original vision of the internet as a space with multilateral publishing and consumption of information, it was meant to be a peer-to-peer vision with no dependency on a singular party, Google has helped take the internet forward both technologically and in adoption, they lead the efforts to improve

[18] J. Clement, *Facebook: number of daily active users worldwide 2011-2019*, April 26, 2019, Statista, Available at:[https://www.statista.com/statistics/346167/facebook-global-dau/]

[19] J. Clement, *Mobile internet usage worldwide - Statistics and Facts*, July 5, 2019, Statista, Available at: [`https://www.statista.com/topics/779/mobile-internet/`]

the open Web, such as via advocating about Progressive Web Apps (PWAs) over native mobile apps. However, their main goal tends towards the accumulation of as much rich, clean and aggregated datasets for their AI backbone. Their mission involves a personalized assistant for each individual user based on their data and requirements in order to provide personalized and timely information to the end user.

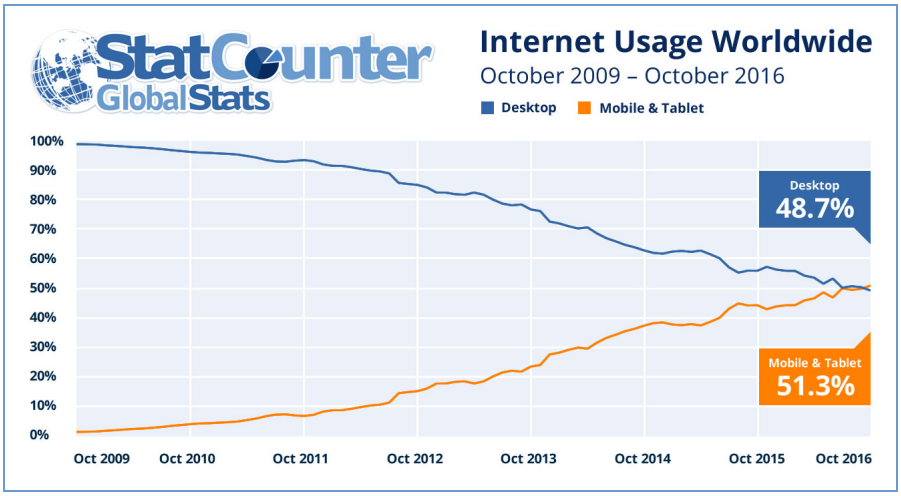

Figure 5: Internet usage worldwide between Desktop and Mobile or other IoT devices (October 2009 - October 2016)

Within the past decade, mobile internet usage has shown astounding growth worldwide due to the advancement of cellular infrastructure making internet accessible to the people even in the most remote areas. Figure 5 depicts the rise in the internet access via Mobile and Tablets over the past decade and the growth of mobile usage to access internet have surpassed desktops in the October of 2016 by astounding 51.3% of the population have access to the internet via mobile phones whereas only 48.7% of the people use desktops to do so.[20] This exponential growth in the amount of data being generated worldwide is an implication of Moore's law.[21] And now, according to IBM[22] they estimate that,

> 90% of the data that exists today was created in the last two years, with around 2.5 quintrillion bytes of data produced each day from almost every sector of the economy.

In foreseeable future, a proliferation of data on consumer demographics, behaviour and attitudes - including health and location data collected via smartphones and the growth of internet enabled devices will be observed and in certain major financial capitals around the world it has been the case that certain data analytic firms offer these data processing services for various clients for an algorithmic insight into the operational performance of their business model. One of the major clients for these data analytic firms are political parties as discussed in subsection 4.2. (SCL Elections Ltd Indian operations.) Some of the 'key trends' in the data market includes, companies realizing the value in the collection, combination and analysis of diverse datasets; as they are prioritizing the ability to connect and analyze previously discrete datasets. It is in this context natural to account for the severe privacy related challenges this technological shift in data generation and processing poses.

[20] J.Clement, *Worldwide digital population as of July 2019*, Statista, Available at: [https://www.statista.com/statistics/617136/digital-population-worldwide/]

[21] Moore's Law is the observation made in 1965 by Gordon Moore, co-founder of Intel, that the number of transistors per square inch on integrated circuits had doubled every year since the integrated circuit was invented. Moore predicted that this trend would continue for the foreseeable future.

[22] R Jacobson (2013), 2.5 quintillion bytes of data created every day. How does CPG & Retail manage it?, Available at: https://www.ibm.com/blogs/insights-on-business/consumer-products/2-5-quintillion-bytes-of-data-created-every-day-how-does-cpg-retail-manage-it/

4 Background Knowledge: Role of political parties in today's digital campaigning ecosystem.

One of the primary factors that got this whole investigation into motion was the level of political interest and investment in the fine tuning of their political campaigns. The whole political campaigning ecosystem long required a regulatory perspective. As one of ICO's objectives, they wanted to estimate who are the entities that operate in the whole digital campaigning ecosystem, what roles and responsibility do they possess, and how personal data is being collected and used. Political parties are key a part of this ecosystem; they are a key capital providers for political consultancy firms. The commissioner Elizabeth Denham, told the Committee in response to the questions raised in House of Commons committee on the subject of disinformation and fake news,

> "Political parties have a lot of data. They purchase data, perhaps from data brokers. They collect it in various ways. So we need to look at political parties as part of this as well, and make sure that they are able to use new digital campaigning techniques to engage voters and, at the same time, comply with the law and be transparent, and give people control about how they use information."[23]

Political parties serve as data controllers in the digital campaigning ecosystem, it is obvious that they are the clients for political advertising model as without them, there exists no market for third party analytic services that support digital political campaigning. It is these political parties that can truly, in the long run can drive an ethical evolution of political campaigning, evidence of which however is absent. Political parties are the drivers in ensuring a high standard of data protection which is in all sense, beneficial for their citizens.

It is important to not undermine the role of political parties in a democratic society - in a parliamentary system, they are responsible for making the voters hold the government to account possible as by their nature, they provide a vital link between the citizen and the state. In their defence, it is crucial to their functionality to be able to benefit from building and accessing databases that cover the entire voters population; this privileged position allows them to establish an effective line of communication with the electorate.

The proliferation of big data and the exponential growth of social media platforms over the last decade have made a substantial change in how political parties use data. The notion of targeted advertising is not new, however big data and social media have made is possible for political parties to use digital advertising techniques to target their voters audience with highly personalized advertisements based mainly on their personal information most of which is sourced from both social media and computing devices - in relation to their interests and lifestyle. In the same context it is understandable that these parties and campaigns want to gain maximum insights by employing these techniques due to how closely elections are fought today. However, to what extent they are willing to go in order to employ these techniques for their self-centered political agenda is matter of grave concern and a question about not just regulatory enforcement but also a moral one.

[23]Oral evidence: Disinformation and 'fake news', HC 363, 27 Nov 2018, House of Commons (Digital, Culture, Media and Sport International Grand Committee), Q4288, Available at:[http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/digital-culture-media-and-sport-committee/disinformation-and-fake-news/oral/92924.html]

4.1 Introduction to ICO's investigation into the Facebook-Cambridge Analytica scandal and their involvement in 2016 US presidential Elections.

The investigation was launched in the May of 2017 which uncovered a stunning series of events that took place during 2017 and 2018. Eighteen months later, according to the report from the Information commissioner's office during July of 2018 titled "Investigation into the use of data analytics in political campaigns",

> "multiple jurisdictions are still struggling to retain fundamental democratic principles in the dace of opaque digital technologies." [24]

Various parliamentarians, journalists, civil societies and citizens have woken up to the fact that transparency is the cornerstone of democracy. Protecting such a fundamental pillar in any democratic republic is essential for its integrity and it should be maintained with highest concern and commitment. We as citizens can only make truly informed choices about who to vote for if they are sure that those decisions have not been unduly influenced by any party or political campaign .
The nature of this investigation being strictly comprehensive and forensic which uncovered a disturbing disregard of voters' personal privacy. Social media platforms, political parties, data brokers and credit reference agencies are now questioning their own processes sending ripples through the big-data ecosystem. The Facebook-CA scandal was a benchmark privacy breach that set the tone to how personal sensitive user data should be treated and processed.

> "This investigation has become the widest investigation of it's type by any Data Protection Authority encompassing online social media platforms, data brokers, analytics firms, academic institutions, political parties and campaign groups." (ICO, July 2018)

4.2 Analysis of Cambridge Analytica's involvement in Trump's campaign.

However, with respect to the scope of this section, The main focus would be on underlining how Cambridge Analytica (a data analytic firm) managed to access unlawfully huge amounts of sensitive user data from Facebook (a social media giant), harvested, processed and then sold to various electoral campaigns specifically 2016 US presidential election. Figure-3 [25] highlights the crucial steps taken by Cambridge Analytica in order to obtain and transform user's likes into sophisticated profiles which were addressed with highly personalized political advertising.
During an interview with the former director of research at Cambridge Analytica, Christopher Wylie stated how the company managed to harvest such a diverse and vast data set of over 87 million users. During an extensive interview with The Guardian, he explained the processes involved in building up to the scandal. He described the first step of building the training set as the most important step, the data for the training set was collected by Dr Kogans application on Facebook's platform. The 120-question personality quiz along with different incentives to fill it up provided Cambridge Analytica the training set they required in order to predict a person's psychological profile.

[24] ICO, July 2018, https://ico.org.uk/media/2259369/democracy-disrupted-110718.pdf

[25] Alex Hern, The Cambridge Analytica Files, Cambridge Analytica: How did it turned clicks into votes, May 2018, The Guardian. Available at: [https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie]

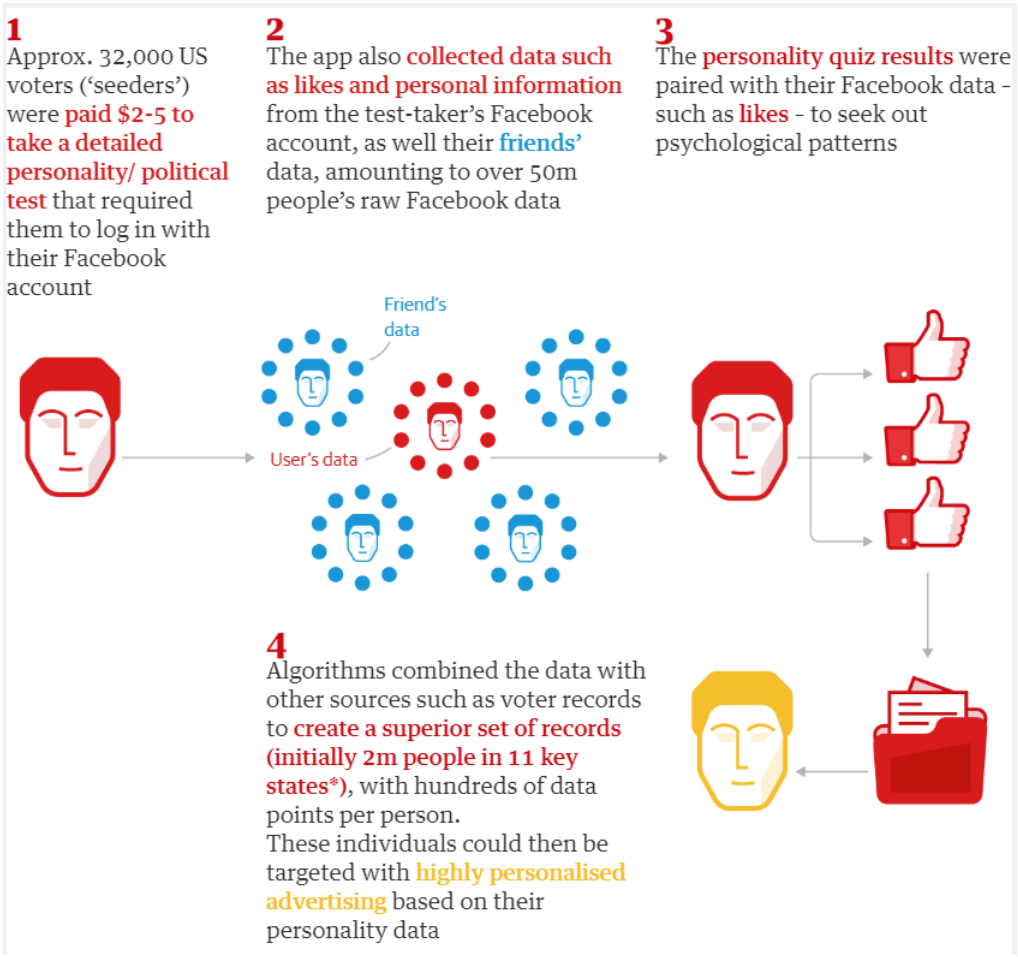

Figure 6: A brief assessment of the Facebook-Cambridge Analytica scandal.

The completeness of the training set was represented by the feature set it contained, that is the underlying data upon which the predictions were made on. In his statement,

> "In this case, it's Facebook data, but it could be, for example, text, like natural language, or it could be clickstream data - the complete record of your browsing activity on the web. Those are all the features that you want to [use to] predict." [26]

Upon assembling your training set, Wylie went on to elaborate further, the next stage in the process was determining what to predict? In this case, it was personality traits and political orientation. Facebook data which lies at the middle of the fallout of Cambridge Analytica proves to be fairly resourceful in the data science community. The personality assessment to profile people were done along five discrete axes in accordance with the OCEAN model. The model clusters personality traits into distinctions that seem to hold across culture and time.

During the testing phase, Wylie stated that Facebook was barely involved. The surveys were offered on commercial data research sites -Amazon's Mechanical Turk platform, then a specialist operator called Qualtrics. It was at the very end when Facebook came into the picture. In order to receive the incentive, which was $2-$5 dollars, for completing the survey the users were required to log into Facebook and approve access to the application developed by Dr Kogan. The process was quick and simple, the user clicks the application and it gives you the payment code. However, two very important things happen in the seconds between the click and receiving the payment code. As Wylie states,

> "First, the app harvested as much data as it could about the user who just logged on. Where the psychological profile is the target variable, the Facebook data is the feature set: the information a data scientist has on everyone else, which they need to use in order to accurately predict the features they really want to know." [27]

[26] ibid, https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christop
[27] ibid

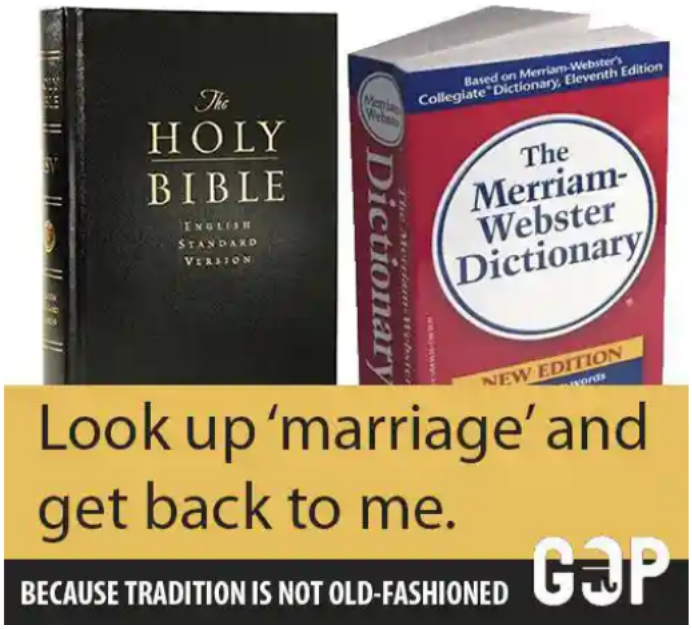

Figure 7: One of the many political advert designed and tested by Cambridge Analytica for Trump's campaign during 2016 US presidential elections

What makes it a major data breach is that this process also provided identifiable information such as real name, location and contact details. Which means, 'that you can take the inventory and relate to a natural person who is matchable to the electoral register.'(ICO, Nov 2019)
In addition to this, the application did the same thing for all the Facebook friends of the user who installed the app and suddenly, hundreds of thousands of people who were paid a couple of dollars, whose personalities were unknown become millions of people whose Facebook profiles are an open book. The final transformation was the transformation of turning a few hundreds thousand personality profiles into a few million, which naturally required a lot of computational resources. At the end of a rigorous ensemble orientated analysis of the data acquired by Cambridge Analytica, they built '253 algorithms, 253 predictions per profiled record.' The objective of Cambridge Analytica was achieved with these models, which was to effectively take the Facebook likes of its subjects and reverse engineering back to their personalities and political affiliations and way more other personality based assessments. As Wylie uncovered the success behind the adverts created by Cambridge Analytica as the best there ever were; he stated in the interview with The Guardian,

> " a neurotic, extroverted and agreeable Democrat could be targeted with a radically different message than an emotionally stable, introverted, intellectual one, each designed to suppress their voting intention"

For instance, one of the adverts designed and tested by Cambridge Analytica for 2016 US presidential is shown in Figure 8, according to Wylie this advertisement was designed to target conscientious people; as for someone who is conscientious this picture of a dictionary reflects a well designed structure and conscientious people are lean more towards order. This message was particularly used in order to boost right wing turnout during the 2016 US elections. As Wylie puts it himself,

> You can create messaging that doesn't make sense to some people but makes so much sense to other people.

Their methods and strategies regarding Trump's campaign were a crucial factor in its success; However, one of the major objectives during the run up to the elections for Trump's campaign was voter suppression. The campaign focused on three main groups which were potential Clinton supporters (white liberals, young women and African Americans) and they achieved so by specifically designing adverts that discourage them from voting.

> "Pretty much every message that Trump put out was data-driven," says Cambridge Analytica CEO Alexander Nix.[28]

Trump's team rigorously tested 175,000 different advert variations based on Donald Trumps argument the testing was conducted in order to find the best combination or individual advertisements worked the best to suppress potential Clinton voters. The testing of these adverts was conducted on Facebook as sponsored news-feed-style adverts that can only be seen by users with specific profiles. The difference between each of these variation existed in the microscopic sense i.e. different headings, colors, captions, with or without any multimedia content. As Nix further explains that these adverts are capable of reaching down to the smallest group of people. According to the Interview with Vice News, Nix stated on the targeting ability of Cambridge Analytica "We can address villages, or apartment blocks in a targeted way. Even individuals."[29] One of the instances that highlights this suppression of voters was on Miami district of Little Haiti where Trumps campaign delivered inhabitants with the news about the failure of the Clinton Foundation following the devastating earthquake in Haiti, with the primary objective to keep potential Clinton voters from the ballot box.[30] It is important to note exactly how precisely these methods were working and more importantly how the American population was being targeted by Trump's campaign as they used personalized messages on social media and digital TV, as it wasn't clearly visible during the time. However, now we have somewhat of a clear picture of Trump's advertising strategy. In a report on presidential advertising by the Political Advertising Resource Centre of University of Maryland; in their analysis of 17 general election ads from Trump's campaign, and 38 ads from Clinton campaign they stated that 2016 campaign ads represented a referendum on candidates' character. Both the campaigns were designed around attacking their opponent through fear and anger appeals. Clinton's campaign used Fear as the primary, underlying appeal in 36.8% of their total ads for the campaign; that means 14 ads out of 38 ads were Fear based, following it the same number of ads were based around Anger. Similar to Clinton's campaign, 10 out of 17 of Trumps ads were coded with either Fear or Anger as the underlying emotion. (PARC, Bhat; Philips; Hess; Hunter; Murphy; Rico; Stephan; Williams, 2016.) This analysis presented a grave conclusion regarding the outcome of 2016 presidential election.

4.3 Presidential Advertising: Analysis of Character attacks through TV advertisement and social media.

Both the campaigns employed Fear and Anger appeals, elevating the tensions, anxiety and cynicism surrounding the elections and this form of mediation has repercussion long after the election. Trump's campaign mainly relied on the long reach of social media, his campaign spent significantly less on TV advertisements, $100 million less that Clinton. Figure-8[31] provides a visual representation on the difference of expenditure by both Trump's and Clinton's campaigns. Trump sincerely relied on character attacks his subject matter for 35% of the TV adverts his campaigns released in the span of 3 months (August till late October 2016.) Out of which 6 were categorized as character attacks. What have been observed in the past two decades is a steady decline regarding the financial investment in TV and radio based ads.[32] One of the factors that seems to be the reason for this drop in spending is due to the expansion of political campaigning to much broader and open platforms of social media and electronic ads. (PARC, 2016) For instance, Clinton's primary and secondary Youtube channel have released over 70 and 116 videos respectively, which are far greater in number than her TV ads which were just 38.
Focusing on Clinton's campaign character based strategies, one of the things that were clearly observed during the analysis of their adverts was the specific use of negative ads to attack Trump's character

[28]Hannes Grassegger and Mikael Krogerus, The Data that turned the world upside down, 28 Jan 2017, Vice News. Available at: `https://www.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win`

[29]ibid

[30]Green and Issenberg, 2016, `https://www.bloomberg.com/news/articles/2016-10-27/inside-the-trump-bunker-with-12-days-to-go`

[31]Adam Pearce, Trump has spent a fraction of what Clinton has on Ads, 21 Oct 2016, The New York Times, Available at: `https://www.nytimes.com/interactive/2016/10/21/us/elections/television-ads.html?_r=0`

[32]Compared to Obama's 2012 campaign's budget for TV based ads which was $378 million dollars.

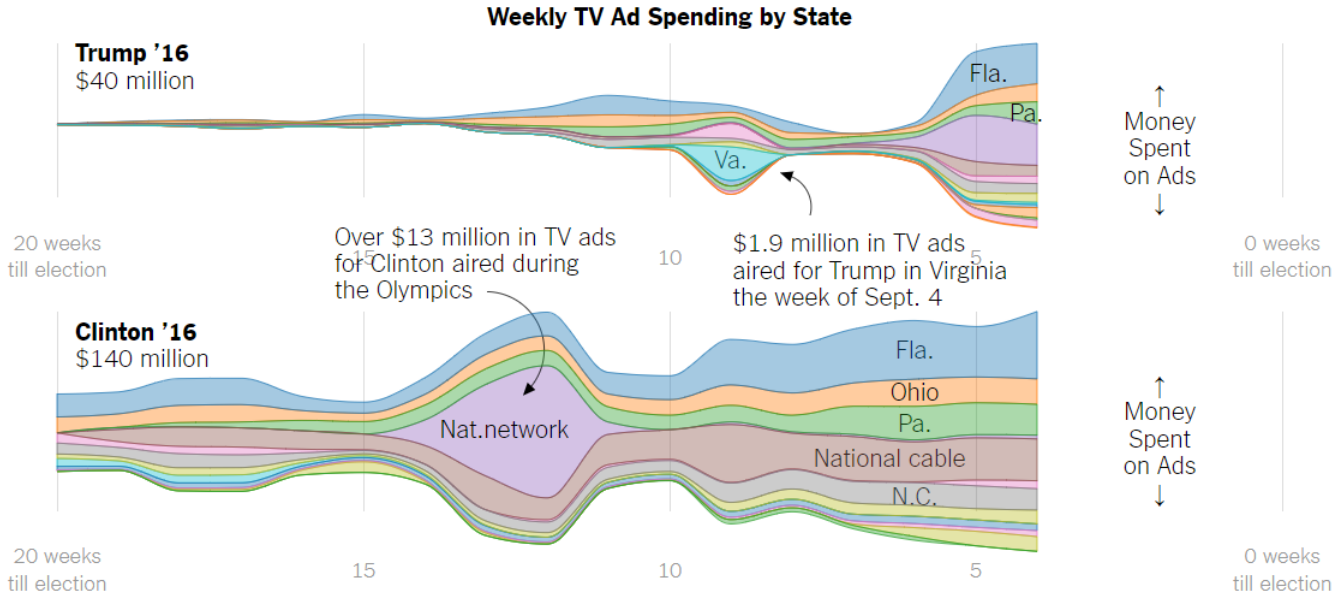

Figure 8: Trump V/S Clinton's campaign expenditure on TV adverts.

instead of his policy positions. Out of her 38 aired adverts, only 10 (26.3%) were predominately focused on policies and issues while the rest of them were character based.

> "One of the way the centrality of character can be explained throughout and after 2016 elections is to understand Donald Trump as an outsider, a candidate who has broken with the customs of decorum as a challenge to how the American political system functions. Trump is a real estate mogul and reality TV star, not a career politician."

Clinton's campaign heavily capitalized on his lack of experience and reality-TV based persona through character attacks, in addition to this, his own words statements about immigrants, person of color, persons with disabilities and women during the campaign allowed turn biggest democratic elections into a referendum on his character. Similar to Clinton, Trump's campaign relied on character attacks as the subject matter for more than 35% of the overall TV adverts released.
Both the campaigns heavily relied on emotional appeals in their advertising strategies in order to shift the public's perception of each candidate's character. Emotional appeals are an effective way to promote democratically desired behavior as they are very effectively coded within the adverts and can change the way of citizen participation in politics and greatly influence their choices[33]. 2016 US presidential elections are the biggest demonstration of these powers at play. While Clinton's ads mostly appealed to the perceptions of joy (17 out of 38 ads) , specifically sought to elicit pride for Clinton, while the second and third most frequently used emotional appeals in her advert strategy was of fear and anger. Adverts such as 'Just one' and 'America's Bully' highlights these emotions. Fear and Anger appeals along with Trump's own controversial behaviour helped frame as a threat to the security and unity of the American people. The main objective of the adverts were to make the citizens perceive Trump's behavior and policies should be met with concern and anger. (PARC, 2016) While Trump's advert strategy was a mixture of positive, negative and comparative. Unlike Clinton's massive advert budget Trump deployed a series of adverts coded with negative emotions such as fear and anger before turning towards more positive messages around the October of 2016. Trump's ad approach was purely reactionary, based on what the situation demanded, it was unpredictable and lacked coherent thought. His "on again, off again," strategy sincerely lacked a discernibly pattern. His primary advert strategy relied on character-based arguments about himself and Clinton. Trump in 10 out of 17 of his ads had the underlying emotion of fear and anger. Through these emotional appeals along with preying on the economic anxieties of citizens, Trump portrayed himself as a true guardian of the American people. In his "Deals" advertisement, the narration follows,"Today our jobs are gone. Factories closed...Donald Trump knows business and he'll fight for the

[33]Ted Brader, "Striking a Responsive Chord: How Political Ads Motivate and Persuade Voters by Appealing to Emotions," American Journal of Political Science 49, no. 2 (2005), 388-405.

American worker." this advert was designed to instigate fear towards the unreliable opponent while, at the same time also instigating anger towards the previous Obama administration. The advert poses Trump as a source of power, success and security - the only one who can improve America by better deals as another one of his main campaign strategy was to strongly emphasize his savvy, success and his ability to reignite the American dream. However, the amount of free media received by Trump along with his uncanny yet effective use of social media gave his campaign the edge it needed against the vast Clinton campaign. By August 2016, Trump had over 22.7 million likes and followers on Facebook, Twitter and Instagram while Clinton had a combined 15 million, now in 2019 alone on Twitter Trump's got 63.6 million followers. What is astonishing and to some extent worrying is that historically, the reach of political adverts have always been proportional to the capital invested by the sponsors and yet, Trump due his proactive and controversial remarks about immigrants and women (mainly), have managed to overcome this. In that sense, political ads had uncharacteristically less relevance to Trump's outreach effort than is typical in most contemporary campaigns.
Through the use of various emotional appeals in political adverts and a well designed character-based campaign across multidimensional news and media outlets, both the campaigns have promoted the message that neither candidate is fit to lead the country. These attacks and political play which have been observed throughout the 2016 elections seems to have changed the conduct of campaigns so drastically that it has made governing more difficult, irrespective of the outcome of the election.

> Since 1952 till 2008 only 31% of the general election ads were character-based, in 2016, character based ads summed up to an astounding 76% of the TV campaign ads.[34]

This open expression of fear and anger towards the opposing campaigns changed the political climate around the elections so drastically that the main issues and problems that need effective and smart resolution dissolve away into petty name calling and cynicism.

5 Background knowledge: Overview of the Legislation (Data Protection Principles and EU's GDPR)

At a glance, data protection is about ensuring people can trust you to use and operate on your data fairly. Strong data protection legislation helps ensuring fair and proper use of information about people, As a part of the fundamental right to privacy to a much more practical level the main objective of data protection regulation is to instate people's trust in organizations that deal with their data. However, along with this data protection regulations are employed to prosecute organizations which breach those regulations.
Data protection is also about getting rid of barriers to trade and help stabilizing co-operation within multinational entities that deal in data. In the form of international treaties for common standards strong data protection regulation enables the free flow of data across borders. Transnational nature of data proposes a significant challenge for national regulatory authorities over the world to enable joint co-operative actions against tech giants like Facebook. Before understanding the principles of data protection we must analyze the definition of personal data and its scope, According to the UK's information commission,

> In short, personal data means information about a particular living individual. This might be anyone, including a customer, client, employee, partner, member, supporter, business contact, public official or member of the public.[35]

There is a clear distinction between private information and personal data as even the information which lies in the public domain or is public knowledge or more so, also the information is regarding someone's professional life can be categorized as personal data whereas private data is hidden from the public domain and remains private to an individual.

[34]PARC, 2016
[35]Some Basic Concepts (Data protection, ICO, Available at:[https://ico.org.uk/for-organisations/guide-to-data-protection/introduction-to-data-protection/some-basic-concepts/])

As long as that information is capable of identifying someone, either alone or by combining it with other information, it is considered as personal data. After UK's adoption of EU's General Data Protection Regulation the definition of personal data now have a wider boundary. Considering the identification of an individual through the information they may or may not have provided as the key element that makes data personal, it is crucial to understand what information can be used and processed in order to identify an individual; it could be as simple as name, phone number, email address or other identifiers such as an IP address or cookie identifier. Now, the information can either directly or indirectly identify an individual. Compared to direct identification indirect identification is quite extensive as it requires to analyze that even after an individual is not directly identified by the information, weather they are still identifiable upon any processing of the information.
The ICO suggests the entities that deal in personal data to take into account the information they are processing along with all the means reasonably likely to be used to identify the data subject(s). However, in relation to weather or not an individual is identifiable or identified either directly or indirectly, from the data or information that has been collected and/or processed, by the DPA 2018 that data or information is not considered as personal data unless it 'relates to' the individual. When it comes to whether the information relates to an individual a wide range of factors comes into play such as, the purpose or purposes for which the information is being taken and processed and its likely impact or effect on the individual.[36] Different purpose for the same information allows it to be personal data in some cases and not in others as for some data controllers that information is considered to be personal data and for some its not personal data. One of the key attributes of GDPR is that only truly anonymous information is not covered under its scope. Information with all identifiers removed or replaced in order to pseudonymise the data is still person data for the purpose of GDPR.
The meaning of whether data or information 'relates to' an individual is more than just simply identifying an individual. The information or data collected or processed must concern them in same way like the content posted on Facebook's platform by 2.7 billion people. In order to decide weather or not data or information relates to an individual the ICO suggests the following to consider,

- the content of the data - is it directly about the individual or their activities?;
- the purpose of which data is being processed for and;
- the results of or effects on the individual from processing the data.[37]

This definition poses a conundrum as there could be data that can refer to an identifiable individual and not be personal data, as the information does not relate to them, such as inferred data used by political parties. In the scope of this section, i will try to differentiate between data controllers and processors as it is crucial to understand the difference between the two in order to understand which GDPR obligation apply to which organization. Data controllers are the primary decision makers as they dictate the purpose and means of processing data, for instance, Facebook and their services can be considered of as data controllers but at times an organization can also act as data processors too, as the purpose of processing the personal data act as a crucial element in the definition of both.
In order to determine whether an organization act as data controllers they must consider the degree of control they have over the purpose of processing personal data. If an organization dictate an overall control in the decision making process of how and why personal data should be processed then that organization acts as a data controller. However, when it comes to being a data processor, an organization which do not have a particular, direct purpose for processing personal data except acting on the behalf of and only on the instructions of data controller. Both the data controller and processors are liable to supervisory authorities such as UK's information commission are bound to act within the scope of data protection legislation. Supervisory authorities keep the practices of data controllers and processors in check as they are capable of taking relevant actions against controllers and processors if they breach the obligation(s) put forward by EU's GDPR. In this context, data controllers have more obligations under the GDPR as compared to the processors.

[36] ibid
[37] What is personal data?, ICO, Available at: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/what-is-personal-data/

5.1 Understanding Data Protection Principles: DPA 1998 and GDPR

Data protection principles in EU's GDPR are laid out at the start of the legislation and inform everything that follows. These principles are crucial for safe, lawful, co-operative and effective business practices for anybody who deals in personal data. They are not rigid in nature as they incorporate instead, the spirit of the GDP regime. At a glace they are stated as seven key principles;*Lawfulness, fairness and transparency, purpose limitation, Data minimization, Accuracy, Storage limitation, Integrity & confidentiality (security) and Accountability.* In the scope of this section, i shall provide a surface analysis of these principles starting with the first data protection principle regarding the lawful, fair and transparent processing of personal data with regards to the data subjects.
This principle strictly implies the simplicity in a manner of consent and communication of personal data as it states about being clear, open and honest with the data subjects from the beginning about the purpose for which their personal data is being used. GDPR dictates six grounds for lawful processing for any entity that deals in personal data, namely, **Consent**, which means that the data subject has willingly agreed for you to process their personal data given the fact that a reason for which their personal data is being provided to them with clear and simple terminology. **Contract** this basis for processing of personal data involves a necessity for a contract that the data controller and the data subject have and it is by the contract that the processing is necessary. The basis of **Legal Obligation** for the processing of personal data states that the processing is necessary for you to comply with the law excluding all other contractual obligations.[38] Another base for lawful processing of personal data incorporates **Vital interest** which states that the processing of personal data is necessary to protect someone's life, along with vital interest, if the processing of personal data is crucial for you to perform a task in the public interest or for official functions them the processing of personal data falls under lawful processing given that the task or function has a clear basis in law. **Legitimate interests** the basis of legitimate interest for lawful processing is a bit less straightforward, it goes on and states that the processing of personal data has a lawful basis if the processing is necessary for your legitimate interests or any third parties, however and unless there is a good reason to protect the individual's personal data which automatically overrides those legitimate interests. (excluding public authorities)

As we observe many lawful bases for processing directly depends on the processing being necessary or not. The ICO in their guide to the general data protection regulation describes that the term "necessary" does not mean absolutely essential but it must be more than just standard practice. The lawful basis only hold when the processing is necessary, they do not apply when the specific objective for which processing is necessary could be achieved by some other less intrusive means or by processing less data. [39] Along with having a lawful basis the handling of personal data must be coherent to the public's reasonable expectations and **not use personal data in ways that have unjustified adverse effects on them.**
The second data protection principle is about **purpose limitation** ensures clarity and openness regarding the collection of personal data. Along with many similarities to the second data protection principle in the 1998 Act, GDPR's principle of purpose limitation differs only in one way, under the GDPR one can comply with purpose limitation principle by complying with their documentation and transparency obligations, rather than by registering with the ICO. GDPR outlines the seriousness of specifying the purpose for processing of personal data as it helps whoever that is dealing with personal data to be accountable for purpose they are processing the data. However, the legislation does not account for future change in purpose for processing of personal data if the purpose complies under public interest, scientific, statistical purposes. This is an explicit exception under the GDPR. **Data minimization** in GDPR is almost identical to the 'adequacy' principle of DPA-1998, the only difference comes under the demonstration purposes as the party dealing with personal data must remain prepared to ensure and demonstrate that their data minimization practices are in line with the new accountability obligations.[40]

[38]Article 6, GDPR, 27 Apr 2016, Available at:[https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1528874672298&uri=CELEX:02016R0679-20160504]

[39](ICO, Guide to GDPR. Available at:[https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/])

[40]ibid, Data Minimization: What's new in GDPR?

The essence of this principle lies in using adequate, relevant and limited to what is necessary amount of personal data in relation to the purposes for which it being processed. This principle proposes a threshold in the form of a minimum amount of personal data one would need to meet their purpose and ensures the usage of only that amount. In addition to using only the minimum amount of data required to meet the purpose for which it being collected for, whosoever is dealing in personal data should also make sure the data do not contain any irrelevant elements as if found so, the liable entity would be in a direct breach of data minimization principle. ICO suggest, whosoever is dealing in personal data to periodically review their processing to ensure the relevance of personal data they acquired and to delete anything that is no longer needed, this enforces the data subject's right to erasure. The fourth data protection principle deals with **Accuracy** of personal data that is being processed. This principle holds whosoever dealing in personal data into account if the data they're processing is incorrect or misleading as to any matter of fact. This implies to consistence updating of personal data relative to the purpose for which it was acquired. The ICO clarifies that one must always be clear about what they intend the record of personal data to show. The principle of accuracy is understandably a difficult and confusing principle, as there is often confusion about whether it is appropriate to keep records of that had happened which should not have happened?
However, there might be cases where one may have to legitimately need the records to very accurately reflect the order of events in which may have occurred. For instance[41],

> A misdiagnosis of a medical condition continues to be held as part of a patient's medical records even after the diagnosis is corrected, because it is relevant for the purpose of explaining treatment given to the patient, or for other health problems.

In cases where the individual challenges the accuracy of their personal data the liable authority must take every measure to consider whether the data is accurate and if not so, it must either be corrected or deleted. It is important in this context to underline the data subjects absolute right to rectification of incorrect data.[42] The fifth data protection principle under the GDPR regime deals with the limitation on storage duration of personal data, it strictly instills that personal data should not be kept for longer that it is required. This is relative to the purpose of collecting and holding the data along with a reasonable justification for how long one might need to keep it. Being very similar to the retention principle of the DPA-1998, the fundamental key remains the same in both the legislations which is, personal data should not be kept for longer than it is necessary. However, under GDPR, whosoever that is dealing in personal data can keep anonymised data for as long as they want. GDPR also do not account for keeping unanonymised data for public interest archiving, scientific or historical or statistical research purpose, given that appropriate safeguards have been employed to minimize the chances of any sort of compliance and/or external breaches. The storage limitation principle under EU's GDPR applies to all types of data, as it is up to the controllers and processors to legitimately specify how long they would be storing the data for. In addition to this, upon strict implementation of this principle of either erasing or anonymising the data, the risks related to data becoming irrelevant, inaccurate or out of data reduces significantly. In order to take the first step in towards the documentational compliance for the principle of storage duration, ICO suggests, whosoever that is dealing in different types personal data to maintain retention schedule list the types of records and/or information which you hold - its purpose - reasonable duration of holding it so; along with retention schedules, big organizations, corporations and universities as good practice should instate smart retention policies to establish standard retention periods for different categories of personal data you hold.[43], However, this depends on the scale at which the organization is processing the personal data. A small corporation undertaking say, low-risk processing may not need a retention policy but it is their responsibility to make sure that they comply with the storage limitation principle by regularly reviewing the data they hold and receive these must be done so in compliance with the documentation requirements. The final data protection principle in

[41] ibid, What about records of mistakes?

[42] See Article 5(1)(c), Article 16 (right to rectification) and Article 17 (right to erasure, Available at: [https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1528874672298&uri=CELEX:02016R0679-20160504])

[43] (ICO, Guide to GDPR(Why do we need retention policy?). Available at:[https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/])

GDPR's Article 5(1) is security related, it embodies **integrity and confidentiality.** It demands the assurance of appropriate security measures to be in place when dealing in personal data. This principle in essence mirrors the 'appropriate technical and organizational measures'' in the previous data protection legislation 1998. GDPR however, goes further into the specifics about the standards of security measures of the processing of personal data. The key difference however between the security principle provided by the DPA-1998 and GDPR's integrity and confidentiality principle is that, what was considered in a broad extent a demonstration of good and best service is now backed up by the law under GDPR, making them legal obligations for data controllers. They now must have appropriate security measures to prevent the personal data under their custody to be accidentally or deliberately compromised. In the connection to the Facebook's breach of personal data by Cambridge Analytica's associate firm GSR, according to the ICO's investigation, Facebook was in the direct breach of data protection principles 1 and 7 by the Data protection 1998 and EU's GDPR.

5.2 Eighth data protection principle: Guidelines on International Transfers.

In relation to the above described data protection principles, a separate provision regarding international transfer of personal data is formulated in the GDPR.[44] The purpose of the eighth data protection principle is to provide safety and proper regulation for the transfer of personal data outside the European Economic Area.[45] DPP-8 outlines the necessity of adequate protection that is required when transfer of personal information takes place at a transnational level, in order to protect the rights and freedoms of data subjects regarding the processing their personal information. However, the eighth data protection principle in its scope consist of a number of exceptions since a transfer of personal data can take place to nations without a proper data protection legislation where the adequacy of appropriate safeguards in ensured otherwise. It states,

> Personal data shall not be transferred to a country or territory outside the European Economic Area unless that country or territory ensures an adequate level of protection for the rights and freedoms of data subjects in relation to the processing of personal data.[46]

According to the legislation, the definition of transfers involves relocation of personal data outside the respected jurisdiction rather than it simply passing through. A transfer of data which before the transfer was held as personal data and will most likely be intended to be held as personal data after it reaches it destination is considered as a transfer of personal data according to the legislation. The adequacy of a data protection legislation of a nation outside the EEA must get the proper designation of adequacy by the European commission which enables the data transfer to meet the requirements of the eighth principle. In addition to adequacy transfers can be made with the explicit consent of the data subject. The notion of online consent in its entirety is not merely just the formal notification of "what an organization is going to do with your data" and the data subject has no real choice but to give their consent to the processing and/or transfer of their personal data. This is not a valid example of consent by the law. Consent by definition are bound to be specific and informed by nature, the data subject must know in simple and understandable terms what they're agreeing to; as far as transfers to their countries are involved the consent message outlined by the controllers should reflect the reasons for such a transfer.

[44] Article 44,45,46: General principle for transfers, Available at:[https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1528874672298&uri=CELEX:02016R0679-20160504&print=true]

[45] EU member states, Iceland, Liechtenstein and Norway.

[46] Information Commissioner's Office. (11-July-2018) "The eighth data protection principle and international data transfers", Available at: https://ico.org.uk/media/for-organisations/documents/1566/international_transfers_legal_guidance.pdf

5.3 Examples of proper consent (Vice news and ICO)

In relation to the transfer of personal data and the risks associated with such a transfer, it is by the law an essential duty of the data controllers to inform the data subjects in simple terms and definitions on how, why and where their data is being used. However, each case is specific to its set of circumstances in a manner of designing a general structure of properly informative consent message which is likely to produce a valid consent, controllers must consider a broad range of factors and risks associated with their practice. For instance, a message which is likely to produce invalid consent would be brief and without the specifics of the processes that is within the data subjects rights to information.

> "By signing below you accept that we can transfer any of the information we keep about you to any country when a business need arises?[47]

By assessing the information above there is no way for a data subject to adequately assess any risk associated with the processing of their personal data and how it may be used.
In the context of this section, since GDPR came into effect, websites and online social media platforms were in a way reminded of the importance informing the data subject about the relevant risks regarding the processing and transfers of their personal data. For instance, the figure below reflects the ICO's effort into setting up an example for properly informing the data subjects in a manner that would most likely generate valid consent. Article 7 of GDPR deals with the conditions controllers must consider in order to get valid consent the data subjects in relation to the processing of their personal data. The article 7 clearly outlines the importance of informing the data subjects regarding any additional purpose to which their data might be used for and that to should be presented in a manner which is clearly distinguishable from other matters, in an intelligible, easily accessible form, using plain and clear language.

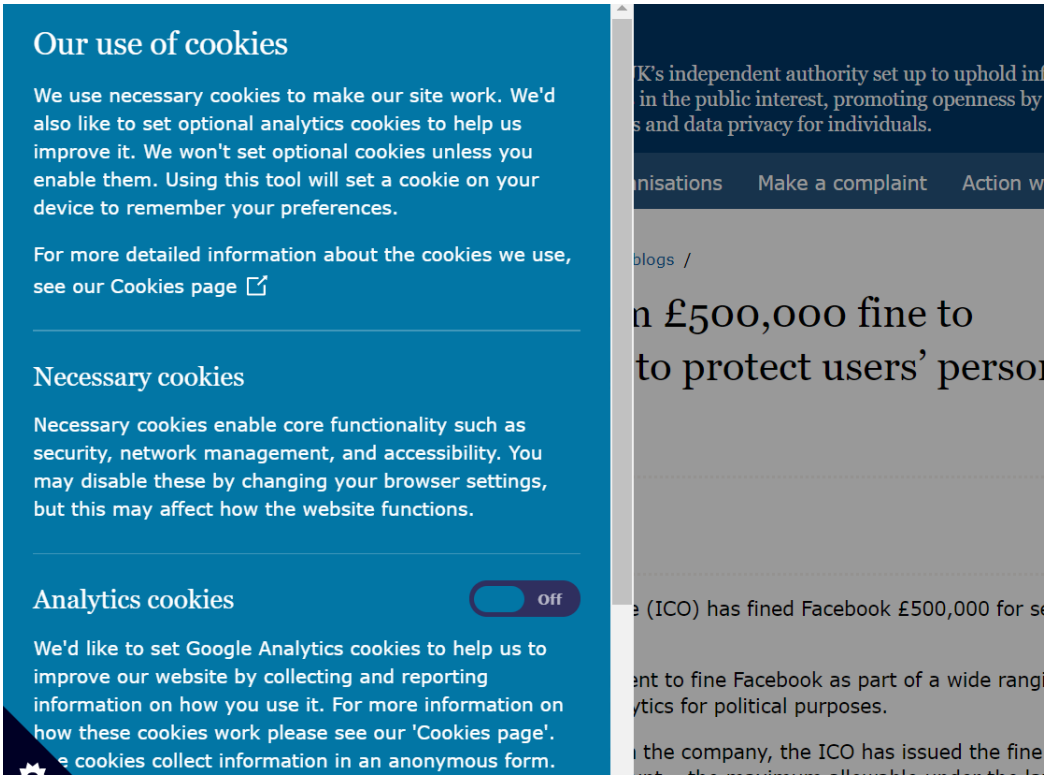

Figure 9: ICO's consent message regarding their cookie policy for website monitoring.

[47] ibid

However, international transfers do not solely depend upon individual consent for this matter. Transfers are also made where certain types of contracts are in place, either legal or whenever there are necessary reasons for substantial public interest.It is in the best interest for the controllers and data subjects if the data subject is clearly made aware for how their data is going to be used and by whom. Figure-10 below is an adequate example for user consent by Vice news, a global multimedia company.

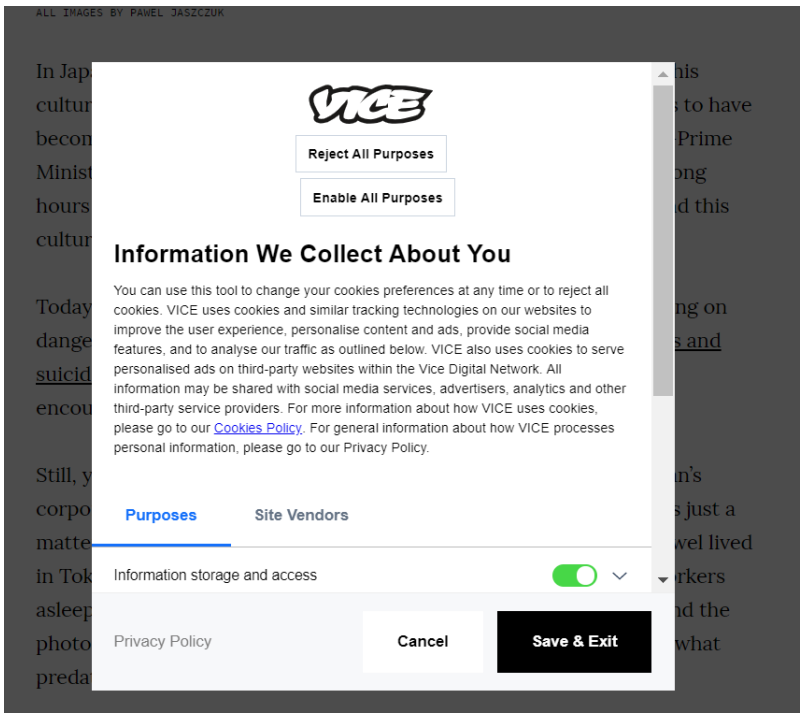

Figure 10: Example of adequate online consent: Vice news.

The clear and simple consent message outlined by Vice news embody all four principles of Article 7 of GDPR that constitutes the conditions for valid consent. The current data protection legislation in the UK is the first step towards a privacy concerned networked society, it covers a broad range of sensitive issues relating to personal data, provides key distinction between processes related to personal data and most importantly, gives an opportunity for controllers and processors to outline their own data protection policies in a manner which is beneficial and safe for the data subjects involved. The section that follows highlights one of the most important feature of the GDPR, its notification power.
Since its effect, for almost a year into GDPR implementation for the first time the European Commission has published infographics regarding compliance and enforcement of the legislation across the EU. The statistics are about the total number of complaints and data breach notification received throughout the first year of the legislations' effect.[48] This transition year have been extremely interesting in terms of observing the GDPR's effect on the EU nations. The data received by the European Data Protection Board shows that the UK's data protection authority is receiving vastly more complaints than other nations, most of the complaints are regarding telemarketing, promotional e-mails and CCTV/Video surveillance. In addition to this, by the GDPR it is mandated that when personal data for which a corporation is responsible is accidentally or unlawfully disclosed, that company is obligated to report the breach to their national data protection authority within the duration of 72 hours after discovering the breach. The figure-12[49] provides a visual contrast among 10 EU nations and the number of complaints submitted to their respective data protection authorities.

GDPR, in its scope is capable of handling breaches notifications throughout the nations it has been adopted by. However, in retrospect of the whole legislation cooperation and consistency mechanism work without any major problems. The EDPB's opinion on the functionality of GDPR underlines the efforts of national data protection authorities to facilitate this cooperation which implies a tremendous amount of workload among national DPAs.

[48]GDPR Today, GDPR in numbers, Open Rights Group, No. 3 25 March 2019, Available at: [https://www.gdprtoday.org/gdpr-in-numbers-4/]
[49]ibid

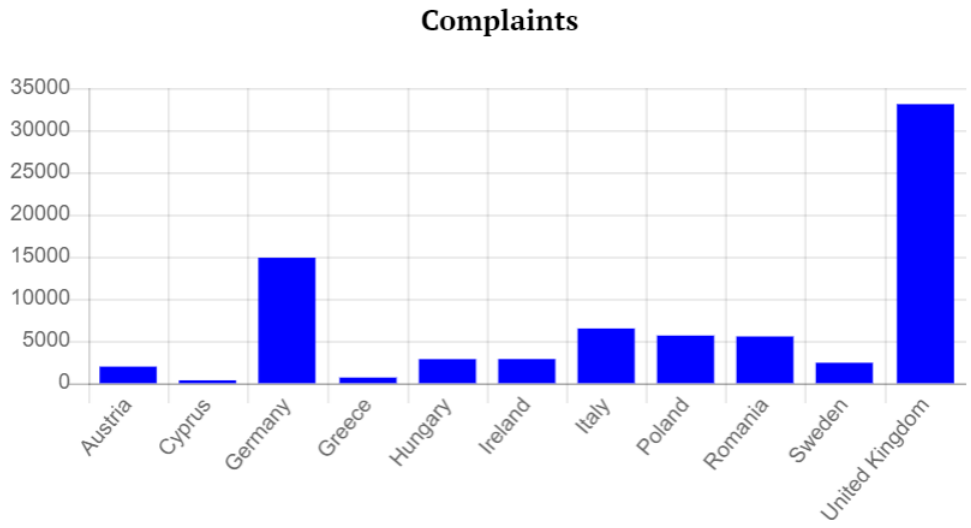

Figure 11: Number of complaints submitted to DPA's of 10 EU nations under GDPR (25 May 2018 to 1 March 2019)

In order to ensure the consistent application of GDPR while dealing with across the border breaches, which require national DPAs to pool in their resources and initiate a joint supervisory action; a general guidance on the interpretation of GDPR seems the best way to take the first step in defining a common interpretation of the data protection legislation and the application of its provisions, by the controllers, processors, supervisory authorities and the public in general.
Now, in the context of data protection legislation both the DPA-1998 and the GDPR the investigation led by the UK's Information Commissioner's office into the matter of Facebook's data breach by Cambridge Analytica after the Guardian surfaced the whole scandal, details of which are elaborated in Annex-1 the data protection breaches made by both Facebook and Cambridge Analytica. The evidence[50] that ICO uncovered during the largest and the most extensive data protection investigation changed the whole cyber world. The most astounding fact from the discoveries made by the ICO is both the entities involved knew to some extent exactly what they were doing.

5.4 Political interest in personal data.

In the UK, after adopting EU's General Data Protection Regulation in effect since May 31 2018; it stands in order to protect personal data of its citizens enforcing rights of the users over their personal data as well as strengthening the enforcement power of data regulators. However, it has been the case that data used by political parties for campaigning purposes is not. 'Inferred' data is not protected; this raises a serious question of distinction between the two categories of user data. According to a report from DCMS on disinformation and 'fake-news', the term inferred data is defined as,

> "the data that includes characteristics that may be inferred about a user not based on specific information they have shared, but via analysis of their data profile."[51]

Inferred data is used as a crucial tool for political consultants and parties to identify their target audience on social media sites, for instance, Facebook provides a means to achieving that through the data profile matching and the 'lookalike audience' targeted advertisement tools. One of the major pillars that this investigation stood on was based on a presumption that political parties do not regard inferred data as personal information, primarily, as it is not factual information. However, if we analyze the implication of inferred data we can see that by definition, it is generated based on the assumptions about individuals interests and preferences. It can also be attributed to a specific individual or a group of individuals.A complete definition allows inferred data to be protected by the data protection law.

[50] Annex-1
[51] DCMS, 2019, Disinformation and fake-news, paragraph 42, Available at:[`https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf`]

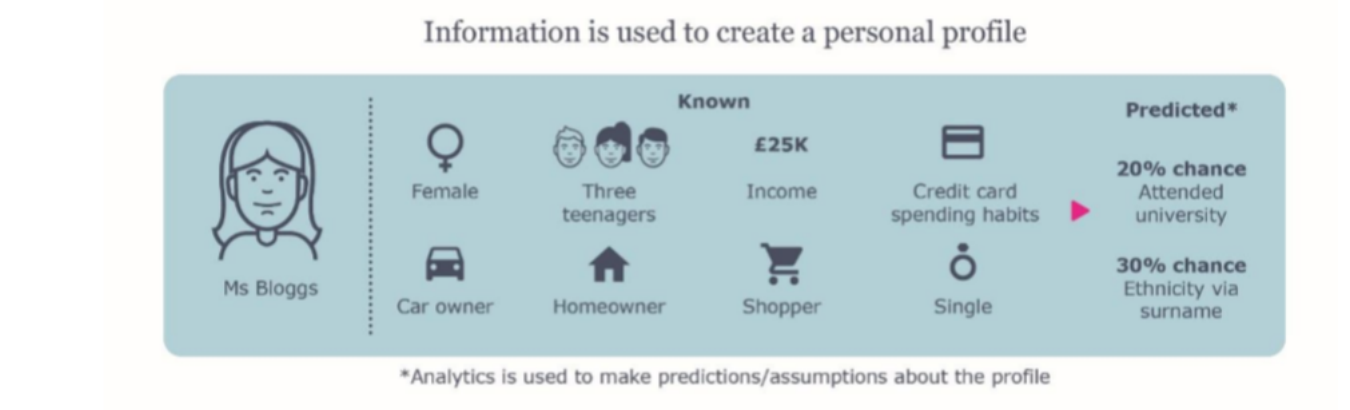

Figure 12: How inferred data can be used to generate a personality profile for an individual.

Figure-12[52] depicts how data analytics is used to generate predictions about an individual using the information related to them, the practice in itself is not new however, due to the datafication of more and more aspect of human life big data analytics now have enough data to make really accurate predictions. The Information Commission regard inferred data as personal data and addresses the problems associated with user awareness and participation regarding data literacy. It is a much difficult issue to resolve when the users are told that they can own their own data, and that they have the power to decide where that goes and what it is used for. The complexity and massiveness of this scandal uncovered a series of flaws in many aspect of handling and safeguarding personal data at an transnational organization while level at the same time highlighting the crucial importance of strong data protection legislation.

> Protecting our data helps us secure the past, but protecting inferences and uses of Artificial Intelligence (AI) is what we will need to protect our future.[53]

The ICO report into the investigation states at numerous accounts how Cambridge Analytica and Facebook deliberately failed to uphold the legitimate, reasonable standards concerning user privacy and data protection. The Information Commissioner herself stated in the House of Commons meeting when asked about the £500,000 fine that was issued by the ICO using the data protection 1998, after being described as a weeks sandwich bill for the tech giant. When asked about the insufficiency of the monetary penalty provided by the DPA-1998, and the need of something more substantial that hits these massive corporations hard in the pocket as been observed in the past from the EU to American institutions in the form of super-fines; the commissioner emphasized on the flexibility in monetary penalties provided by the GDPR,

> "The £500,000 fine that we issued against Facebook was the maximum fine available to us in the previous regime. Because the contraventions happened before 25 May 2018, £500,000 was the maximum. I have been clear that had the maximum fines now available under the GDPR been available to us, we would have issued a much more significant fine to Facebook. We now have up to 4% of global turnover, as do the other EU data protection regulators. That is a significant fine for a company, and it would have an impact on its bottom line. I also think that our findings and the fine that we issued are important, because we found their business practices and the way applications interact with data on the platform to have contravened data protection law. That is a big statement and a big finding."[54]

Political parties are recognised as they key entity operating the digital campaigning ecosystem. The data protection law does not prevent them from using social media in campaign communication nor does the

[52]Information Commissioner's Office. (11-July-2018) "Democracy Disrupted? Personal information and political influence", Available at: [https://ico.org.uk/media/2259369/democracy-disrupted-110718.pdf]

[53]DCMS, 2019, Paragraph 44, Available at: [https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf

[54]Oral evidence: Disinformation and 'fake news', HC 363, 27 Nov 2018, House of Commons (Digital, Culture, Media and Sport International Grand Committee), Q4293-4398, Available at:[http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/digital-culture-media-and-sport-committee/disinformation-and-fake-news/oral/92924.html]

access to the electoral register; allowing them to carry processing in relation to political opinions that their data embody. However, political parties are not exempt from the data protection law either. Their responsibilities as controller is primarily to comply with all the data protection principles. The role of the Information commission is to consider the possibility that whether the collection and use of personal information complies with the data protection law.

5.5 Regulatory issues and Monetary penalties: ICO to Facebook, Cambridge Analytica, SCLE Ltd.

On the basis of evidence that ICO received and recovered, concerns regarding the misuse of data accessed by Dr Kogan and his company GSR from Facebook. The data was then used for the purpose different than its initial purpose of research for which the proper consent of the users were not taken. As per the update by the ICO into the investigation, they are currently investigating whether and to what extent Dr Kogan and other executives of both GSR and Cambridge Analytica are individually culpable in this regard.
Even after numerous of notifications to Dr Kogan and other executives of GRS for voluntary interviews under caution, to provide with their side of the story they have deliberately refused to do so. Following these events, during May 2018 both Cambridge Analytica and SCLE was taken into administration. Since then Cambridge Analytica and SCLE as part of the SCLE Group have ceased trading. The intention of ICO regarding SCLE was of strict and heavy action for the unlawful processing of personal data for political purposes in the 2016 US Presidential campaigns, along with serious breaches of the first data protection principle of DPA-1998.
One of the major objectives concerning the investigation was the processing of data took place in the UK by a UK most of the data in the scandal surrounds the 2016 US presidential elections. Facebook users along with their friends, were not aware of this process and neither were they made aware about it. Outlining the process that their was ultimately used for,

- their data to be transferred to a data analytic firm Cambridge Analytica.
- their personal data would be used for political purposes, targeted political messaging by political campaigns.
- their personal data would be processed in a manner that involved drawing inferences about their political opinions, preferences and voting behaviour.[55]

The ICO's investigation uncovered major events of unlawful processing by the data analytics firm Cambridge Analytica satisfying schedule 2 and 3 of DPA-1998. Neither was their proper consent taken nor the users were made aware in any sense that their data was going to used for political purposes. The nature of the data being sensitive as it was used to make predictions about the subjects' political affiliations and opinions made the nature of the breach more severe as it undermines UK's data protection legislation. The distress among the public is regarding the processing of their data for political profiling without their explicit consent as the breach affected 80 million people world wide, 37 million in north and central America and 1 million in the UK. The objective of the monetary penalties issued by the ICO to get these firms and social media giants to act in compliance with data protection legislation; taking the people they provide services to with them. In order to achieve this compliance and best practice, organizations and regulators must come to general terms and set a standard of practice regarding personal data.
Shortcomings in Cambridge Analytica's practice in regards to data protection law and their IT infrastructure have been highlighted due to this investigation. Among plenty of their failures relating to dealing properly with subject access request one that stands out the most as it paved way for the ICO to proceed with criminal offences against Cambridge Analytica given the gravity of the issue and tremendous public interest concern the investigation raised, is a request submitted in January 2017 by a US based academic

[55](ICO, Nov 2018, P 3.2.3, Available at[https://ico.org.uk/media/action-weve-taken/2260271/investigation-into-the-use-of-data-analytics-in-political-campaigns-final-20181105.pdf)

Prof. David Carroll.
The dialogue between ICO and CA follows,

- Cambridge Analytica initially denied ICO's jurisdiction and Professor Carroll's right to information. Failing completely to respond to the ICO's concerns.
- ICO then served an enforcement notice on 4 May 2018, ordering Cambridge Analytica to comply with the terms of the Subject Access Request filed by Professor Carroll under Data Protection Act-1998 holding them liable to provide with all the copies of personal data Cambridge Analytica held relating to him. along with an explanation as to where they got the data from. The last date to comply with the notice was 3 June 2018
- Cambridge Analytica failed to comply with the enforcement notice issued by the ICO to their denial of service in regards to Professor Carroll's request.
- Criminal proceeding began on 3 October 2018 against Cambridge Analytica by the ICO.
- As the firm pleaded not guilty the trail then was set for 9 January 2019 at Hendon Magistrates Court.

In addition of their non-complaint action, during the investigation ICO seized a number of servers under the warrant of premises search which revealed a messy IT infrastructure. Not only they failed to make sure that the information provided to them by Dr Kogan's company GSR was transferred securely between the company and third parties; usage of personal email addresses and accounts added more security concerns. Along with 'post-it' notes were found on the boards of Cambridge Analytica offices containing passwords. Cambridge Analytica not only failed yet again to delete all the Facebook data in a timely manner, the data might also have been transferred to other bodies and its fragments or copies still very well may be in their possession. (ICO, Nov 2018)
Facebook was issued a monetary penalty by the ICO of £500,000 as the highest fine permitted by the UK DPA-1998 during the Oct 2018. This fine in no way reflective of the severity of Facebook's shortcomings and reluctant nature to change their outlook towards data protection and not to mask it by advertising their concerns towards privacy; neither does this penalty affect the social media giant financially. However, after the whole breach surfaced Facebook's shares plunged 19% in New York, a day after the company revealed that 3 million users in Europe had left the social network since the Observer reported on the Cambridge-Analytica scandal. According to Facebook's chief financial officer, David Wehner,

> Our total revenue growth rates will continue to decelerate in the second half of 2018, and we expect our revenue growth rates to decline by high single-digit percentages from prior quarters sequentially in both Quater 3 and 4.[56]

The number of daily and monthly active users in their major advertising markets such as the EU fell for the first time, that is the number of daily users reduced from 282 million users to 279 million and the number of monthly active users, from 377 million to 376 million. However, in the US the daily and monthly users remained stagnant i.e. 185 million daily active users and 241 million monthly active users. Throughout the plunge, Facebook made efforts in order to highlight that they company's future growth would not initiate from core Facebook's platform but from its other extensions such as their messaging applications, messenger, WhatsApp and Instagram. The CEO, Mark Zuckerberg's vision while including InstagramTV has been an amazing success according to him, in regards to its money-making potential due to its longform video format. Even after lying in the middle of the biggest data protection scandal and investigation ever conducted, Facebook still managed to deliver a 42% year-on-year increase estimated to be around $13.2 billion. The ICO have referred their ongoing concerns about Facebook's targeting functions and techniques that are employed to monitor users' browsing habits, interactions and behaviour across the internet and different devices. (ICO, Nov 2018)

[56]Olivia Solon, Facebook stocks plummet more than 20% amid concerns over growth, The Guardian, 25 Jun 2018, Available at: [https://www.theguardian.com/technology/2018/jul/25/facebook-stocks-second-quarter-revenue-user-growth]

6 Future of Political campaigning

The final section of this study is an analysis of current and emerging trends in how data is being employed in political campaigning. By looking at the current state of the art tends in data itself along with the advancements in the technologies in relation to online political advertising.
Big data, AI and Machine Learning is not something that is new, now a days from smart home appliances to diagnosing diseases the fuel that propels these advances is big data. Vast and disparate datasets that are in a constant state of being added into one another have major implications when it comes to individual privacy rights. The key ingredient that makes they long extensive datasets is often personal data; from the statistics generated by fitness trackers after a run or along walk, to the information generated by the sensors while walking into the local shopping centre to a social media post, the list is enormous and very real. As observed throughout this analysis the vital importance of strong data protection law under which strict rules are applied in relation to the collection and processing of personal data. It is of upmost importance that this discussion is not about "data protection V/S big data." This outlook on these issues is not only wrong but to all extent polarising. Privacy is not a means to an end but in its all entirety it is an enabling right.
In order to understand the implications of big data, AI and machine learning into the practice of political campaigning, we must define objectively and separately each of the three terms. The most common definition of big data as provided by the Gartner IT glossary is,

> High volume, high velocity and high-variety information assets that demands cost effective, innovative forms of information processing for enhanced insight and decision making.[57]

The 3-Vs that give big data its power can be defined in terms of its size that relates to massiveness of datasets accumulated through various different sources where the analysis of such huge and varied amounts of data can take place in real time. The main reason for defining the term is to agree upon its broadest form of definition as there are multiple forms of big data that not necessarily share the same attributes. In regards to Artificial intelligence, the definition provided by the Government Office for Science's constitutes,

> The analysis of data to model some aspect of the real world. Inferences form these models are then used to predict and anticipate possible future events.[58]

In contrast to the classical methods of data analysis, the key difference in AI lies in the non-linear approach of data analysis as AI programs learn form the data they are subjected to in order to provide intelligent responses to new data and adapt their outputs accordingly. AI incorporates giving computers behaviours which would be thought intelligent in human beings.
Big data is a difficult asset to exploit and since the past decade or so more and more use of AI is a direct implication that AI has power to generate value from enormous amounts of varied data. Whereas machine learning is considered as a branch of computer science that underpins and facilitates AI. All together these three disciplines form big data analytics. It is of no surprise that both major and minor business and corporations of both public and private sectors are leaning towards the benefits that the incorporation of big data analytics brings, which is primarily driven by the data generated from new sources such as IoT devices, ultra interactive online social platforms and the digitization of the society in general. However, mostly all current AI systems have a narrow application as they are specifically designed to take a well defined task in a single domain; for instance, the recruitment process in companies employ tools to distinguish obviously unfit candidates. While machine learning remains an exciting academic field and in recent years have seen exponential growth and potential for more, some of the most important breakthroughs in computer science have been achieved via the application of deep learning which consist of the ability to draw meaningful knowledge from large, diverse datasets and make inferences from seemingly innocuous data. For instance, a paper published by the MIT employed deep learning to accurately infer an individual's age and gender based on the metadata concerning their phone

[57] Gartner IT glossary Big data. Available at:[http://www.gartner.com/it-glossary/big-data]
[58] Government Office for Science. Artificial intelligence: opportunities and implications for the future of decision making.

calls; time of call and duration.[59] Along with more and more sophistication in AI, the IoT devices are also experiencing a technological shift as they just might entirely change they way we engage with technology. However, the broad trend is towards even more granular audience segmentation. Facebook has a number of tools which advertisers can employ to reach their targeted audience, one of which is "lookalike audiences" which enables advertisers, political campaigns, whoever who wants to conduct business and reach out to new groups of people who would likely to be interested in their services. In comparison to Google's similar tools such as "customer match targeting." Facebook is keen to further improving the ability to tailor these tools.

> In 2017, Facebook added 'value-based lookalike audiences' for commercial businesses which, according to the site 'creates an additional weighted signal for people most likely to make a purchase after seeing your ad. [60]

These developments are only the beginning to commercialise politics as they are still in their infancy but it would not take a long time before they are integrated to a huge extent into the political landscape of nations who lack adequate data protection legislation and enforcement. For instance, L2 us a data analytics company specialising in data enhancement, offering voter file enhancements where all records in the national voters files are passed through a number of processing stages. Two major aspects of L2 include lifestyle data enhancement and modelling enhancements.[61] In addition to this, L2 already offers a detailed data a voter's perception on divisive issues like immigration, gun control, environmental concerns and so on.

> By combining social media posts with heating bills with health data from a fitness tracker, for example, it might be possible to pick out each users' likely political frustrations or aspirations, then use that to inform the content of adverts.[62]

In addition to this, the rigorous use of social media and commercial advertising techniques has helped change the discourse of the conduct of political campaigning, due to better and better tools for audience segmentation, voter's psychographic analysis allowed campaigns to select and propagate messages on the basis of their response instead of its ideological or political selection.
In the past two or three years, the advancements in AI generated content and natural language processing/generation have set the tone to how effectively AI generated content can be used to spread fakenews and disinformation. Recent development in AI generated content are likely to be used to mislead and confuse using photo-realistic images, imitation of real voices and hallucinated faces have immense potential to promote misinformation in terms of false-proofs that the opposing candidate has done or said something wildly scandalous. The consistence increase in the ownership of IoT devices (specifically wearables) combined with the location data is still in its early stages of useablity in the commercial sector and it is highly likely that this technology can be employed by political campaigns in a manner that gives the campaign strategists new and valuable insights to 'where and when' hold their political rallies and address the general public.

[59]Felbo et al (2017), 'Modeling the Temporal Nature of Human Behavior for Demographics Prediction' Available at:[http://ecmlpkdd2017.ijs.si/papers/paperID90.pdf]

[60]P Robles (2017), 'Facebook adds value optimization to ad bidding Lookalike Audiences', Available from: https://econsultancy.com/

[61]L2 Political ,https://l2political.com/

[62]Information Commissioner's Office. (11-July-2018) "Investigation into the use of data analytics in political campaigns", Available at:[https://ico.org.uk/media/action-weve-taken/2259371/investigation-into-data-analytics-for-political-purposes-update.pdf]

7 Conclusion and Key challenges for future research

In conclusion of this study, i would like to reflect back to the initial question raised regarding the attack on the transparency in the democratic campaigning. This is a rather complex and rapidly evolving area of activity which is elusive in its growth and vague in the level of awareness among the public about how big data analytics functions and how their personal data and information is collected and processed. The analysis of empirical evidences which surface due to the investigations into these matter seems to be the only source to address the questions regarding the disruption of democratic practices by political campaigning in recent years. Regulatory authorities in their scope lack the adequate workforce to address each and every act of misconduct performed by the wide range of entities that operate in the digital campaigning ecosystem. These entities such as social media giants, powerful political campaigns, data brokers and processors and multinational data analytics firms are immensely resourceful. Facebook's internal business model seems to be in a direct conflict with right to privacy for their data subjects as one breach after another, for instance during the April of 2019 another massive breach relating to Facebook was surfaced by the cyber security researches from UpGuard stated that,

> they found two massive troves of exposed Facebook user data that had been posted publicly on Amazon cloud servers. The data included users' passwords, names, comments, and likes. The scope of this particular privacy foul from Facebook is tremendous: More than 540 million user records were sitting in plain sight, available to anyone who found them.[63]

These breaches are a direct repercussion of the severe lack of strong data protection laws in the US, there is currently no comprehensive federal data privacy law in the United States or even a federal requirement for companies to notify users if they have been swept up in a data breach or other violation of privacy wherein their data was improperly handled. These breaches will keep on happening and it is of no surprise. From 87 million user's data taken and fed into the Trump's 2016 presidential campaigns to 540 million users personal data posted publicly on Amazon's servers, Facebook's attitude towards their user's privacy is clearly depicted by their seriousness towards addressing these major faults in the corporation, they are in no rush to put a halt to these series of breaches.
Transparency in the processing of personal data is an key element to fairness and fair processing of personal data establishes and strengthens the public trust in corporations that process and handle their personal data. Lack of trust due to unfair processing and opacity in the processes involving personal data is harmful for both the general public, corporations and severely undermines democratic practices. Nothing about the 2016 US presidential elections was transparent.

One of the many key challenges relating to the future of political campaigning would be, how to deal with the potential for AI generated content as this technology is getting more and more sophisticated and an integral part of political campaigning in the coming years. One of the possible cases would involve personally targeting each individual potential voter with political content specifically tailored for them using their own personal data, held by political parties. This would not only result in inappropriate, misleading and prejudicial advertising appeals; this demeanor of political advertising would further weaken the public's trust in political parties. In the near future the deployment of numerous algorithmic techniques will soon be a part of nearly every political campaign, enabling them to run and send out millions of political messages and adverts the scale of which this will happen will for certain overwhelm data protection authorities. Within the next decade the role of data protection authorities would be of upmost importance as they would require to collaborate and cooperate with other DPAs to deal adequately the situation that is already out of hand.

[63] April Glaser, Another 540 Million Facebook Users' Data Has Been Exposed, 3 April 2019, Slate, Available at:[https://slate.com/technology/2019/04/facebook-data-breach-540-million-users-privacy.html]

8 Bibliography

[1] Information Commissioner's Office. (11-July-2018) "Investigation into the use of data analytics in political campaigns", Available at:[https://ico.org.uk/media/action-weve-taken/2259371/investigation-into-data-analytics-for-political-purposes-update.pdf]

[2] Information Commissioner's Office. (11-July-2018) "Democracy Disrupted? Personal information and political influence", Available at: [https://ico.org.uk/media/2259369/democracy-disrupted-110718.pdf]

[3] Information Commissioner's Office. (11-July-2018) "The eighth data protection principle and international data transfers", Available at: [https://ico.org.uk/media/for-organisations/documents/1566/international_transfers_legal_guidance.pdf]

[4] DEMOS (10-July-2018) "The Future of Political Campaigning", Available at: [https://demosuk.wpengine.com/wp-content/uploads/2018/07/The-Future-of-Political-Campaigning.pdf]

[5] Information Commissioner's Office. (11-July-2018) (06-Nov-2018) "Investigation into the use of data analytics in political campaigns", Available at: [https://demosuk.wpengine.com/wp-content/uploads/2018/07/The-Future-of-Political-Campaigning.pdf]

[6] PNAS, Michal Kosinski, David Stillwell, and Thore Graepel, Private traits and attributes are predictable from digital records of human behavior (09-Apr-2013) Available at: [https://www.pnas.org/content/110/15/5802]

[7] Jenny Rosenberg, Nichole Egbert, Online Impression Management: Personality Traits and Concerns for Secondary Goals as Predictors of Self-Presentation Tactics on Facebook, Journal of Computer-Mediated Communication, Volume 17, Issue 1, 1 October 2011, Pages 1–18, Available at: [https://doi.org/10.1111/j.1083-6101.2011.01560.x]

[8] BARNEY, D. D. (2004). The network society. Cambridge, UK, Polity.

[9] Manuel Castells. 1997. The Power of Identity. Blackwell Publishers, Inc., Cambridge, MA, USA.

[10] Vergeer, M. (2013). "Politics, elections and online campaigning: Past, present and a peek into the future." New Media & Society, 15(1), 9–17. Available at: [https://doi.org/10.1177/1461444812457327]

[11] Information Commissioner's Office. (22 May 2019), "Guide to the General Data Protection Regulation (GDPR)", Available at: [https://ico.org.uk/media/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr-1-0.pdf]

[12] Compaine, B., and D. Gomery (eds) 2000. Who owns the media? Competition and Concentration in the mass media industry. Mahwah, NJ: Lawrence Erlbaum.

[13] McChesney, R., E.M. Wood and J. B. Foster (eds) 1992. *Capitalism and the Information Age: The Political Economy of the Global Communication Revolution.* New York: Monthly Review Press.

[14] Norris, P. (2001). *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide.* Cambridge: Cambridge University Press.

[15] European Data Protection Supervisor (19 March 2018). EDPS Opinion on online manipulation and personal data Available at: [https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf]

[16] Information Commissioner's Office. 2016, Big data, artificial intelligence, machine learning and data protection. Available at: [https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf]

[17] House of Commons Digital, Culture, Media and Sports Committee, Eighth Report of Session 2017-2019, Disinformation and 'fake news': Final Report, Available at: [https://publications.parliament.uk/pa/cm201719/cmselect/cmcumeds/1791/1791.pdf]

[18] Information Commissioner's Office, (6 Nov 2018), Investigation into the use of data analytics in political campaigns, Available at: [https://ico.org.uk/media/action-weve-taken/2260271/investigation-into-the-use-of-data-analytics-in-political-campaigns-final-20181105.pdf]

[19] Oral evidence: Disinformation and 'fake news', HC 363, 27 Nov 2018, House of Commons (Digital, Culture, Media and Sport International Grand Committee), Q4293-4394, Available at:[http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/digital-culture-media-and-sport-committee/disinformation-and-fake-news/oral/92924.html]

[20] ICO guidance, Guidance on political campaigning, Available at: [https://ico.org.uk/media/for-organisations/documents/1589/promotion_of_a_political_party.pdf]

[21] Grassegger, H. and Krogerus, M. (2017). The Data That Turned the World Upside Down. Motherboard [online] Available at: https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win

[22] Ted Brader, "Striking a Responsive Chord: How Political Ads Motivate and Persuade Voters by Appealing to Emotions," American Journal of Political Science 49, no. 2 (2005), 388-405.

[23] Bhat; Philips; Hess; Hunter; Murphy; Rico;Stephan; Williams, A Report on Presidential Advertising and the 2016 General Election: A Referendum on Character, The Political Advertising Resource Center Center for Political Communication and Civic Leadership Department of Communication/College of Arts and Humanities University of Maryland, PARC 2016, Available at: [https://parcumd.files.wordpress.com/2016/11/parc-report-2016-v-21.pdf]

[24] Shepard, Steven. "Trump's Bizarre Ad Strategy." POLITICO. Available at: [http://politi.co/2cnk9KI.]

[25] Information Commissioner's Office, ICO opening remarks - The Committee on Civil Liberties, Justice and Home Affairs (LIBE) of the European Parliament – Hearing on the Facebook/Cambridge Analytica case, 04 June 2018, Available at:[https://ico.org.uk/media/about-the-ico/documents/2259093/ico-opening-remarks-ep-libe-facebook-cambridge-analytica-20180604.pdf]

[26] McDermott, Y. (2017). Conceptualising the right to data protection in an era of Big Data. Big Data Society. https://doi.org/10.1177/2053951716686994

[27] Care Quality Commission (UK), Cookies: Information, Available at:[https://www.cqc.org.uk/about-us/our-policies/cookies#sm-accord-1]

[28] GDPR Today, GDPR in numbers, Open Rights Group, No. 3 25 March 2019, Available at: [https://www.gdprtoday.org/gdpr-in-numbers-4/]

[29] European Data Protection Board, First overview on the implementation of the GDPR and the roles and means of the national supervisory authorities, Available at: [http://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/LIBE/DV/2019/02-25/9_EDPB_report_EN.pdf]

[30] P Robles (2017), 'Facebook adds value optimization to ad bidding Lookalike Audiences', Available from: https://econsultancy.com/

[31] Felbo et al (2017), 'Modeling the Temporal Nature of Human Behavior for Demographics Prediction' Available at:[http://ecmlpkdd2017.ijs.si/papers/paperID90.pdf]

A Crucial discoveries uncovered by ICO's investigation into the Facebook-Cambridge Analytica scandal.

In order to fathom the importance of this investigation, its important here to briefly look upon the scale of the data breach. It is spectacularly massive in nature and depth. The highlight being, personal data of over 87 million users globally, out of which 1 million users in the UK was collected and processed without their proper concern. The investigation uncovered that an application, **thisisyourdigitallife** developed by Dr. Aleksandr Kogan and his company Global Science Research (GSR), harvested the data, some of which was then used by Cambridge Analytica, in order to target the voters during the 2016 US Presidential campaign process.

> "A number of companies including Cambridge Analytica Limited (UK), SCLE Elections Limited (SCLE) operated as a part of the SCLE Group of Companies (SCLE) under the more publicly familiar trading name Cambridge Analytica (CA)." (ICO, July 2018)

Cambridge Analytica was an umbrella company of a parent firm known as SCL election Ltd before its dissolution in 2018. SCL election Ltd a political consultancy firm, which has a clientele ranging from various governments and militaries. Their services primarily involves consultancy for political campaigning. Were they involved in this breach either explicitly and/or implicitly based on the factual evidences uncovered by the investigation and used throughout this thesis is up to the reader to decide however, denying any correlation between a certain degree of political interest over the resourcefulness of data analytics seems willful ignorance. During this whole turmoil the central focus has been on Cambridge Analytica which is an understandable turnout, also contravened by the law. However, the foundations of the targeting techniques and their development which served at the heart of this issue dated back to early 2006 at the Psychometric Centre of Cambridge University.
During early 2008, Facebook deployed Version-1 of their Graph Application Platform Interface (API), which allowed third party applications such as the one developed by Dr. Kogan, to access the huge pool of sensitive data concerning Facebook users and the people in their friends list. To obtain this information the application developers had to request permission directly from the application users prior to its use which was buried deep within Facebook's usage policy or their terms and conditions; this authorization from the users allowed the developers to access the user's as well as their Facebook friends personal information.

Despite of producing and deploying a range of platform policies for developers who deployed applications on Facebook's platform they still managed to collect 700 terabytes of personal data, as a result of the investigation, ICO concluded based on a number of factual evidences that Facebook failed to take sufficient steps to prevent applications from collecting data in contravention of data protection law. During 2014, Facebook made a major change in their platform policy by updating the Version-1 of their graph API to Version-2 and started transferring third party apps from API VI to V2 . Any new third party application then deployed were automatically added to the Version 2 of the API and did not have access to Facebook friend data. During the years 2014-15 when Facebook was making changes to its platform API, their policies for the stated period permitted third-party applications to access and obtain personal data about the user who installed Dr Kogan's application, and in certain cases, the data to user's Facebook friends. According to paragraph (3.2.1) from the ICO's report into the investigation,

> "However, Facebook's platform policy sought to impose limitations on what this data could be used for – it was focused on providing for enhanced user experiences, and did not extend to its use for commercial purposes. Any terms of service changes used by app developers were supposed to comply with Facebook's terms of service and policies, and developers should have been aware of this." (ICO, Nov 2018)

However, during 2013, the Psychometric Center at Cambridge university was carrying out work on psychometric testing.

A.1 Overview of Psychometric Testing Center at Cambridge University.

Dr. David Stillwell and Dr. Kogan continued to develop a number of applications including an app called My Personality based on OCEAN model (The model identified personality traits based on Openness, Conscientiousness, Extroversion, Agreeableness and Neuroticism.) developed in the 1980s. The Academics at the Psychometric centre pioneered the use of Facebook data in connection with the OCEAN model for psychometric testing through the development of the above mentioned online quiz (My personality). The results were calculated in terms of OCEAN scores and these scores were matched with other sorts of online data such as 'likes', 'shares', 'posts' on Facebook in order to develop personality profiles.

> "By referring to as little as 68 likes academics were able to make highly accurate predictions a number of characteristics and traits such as ethnicity and political affiliation." (Kosinski, Stillwell, Graepel, PNAS 2013)

During the same time during 2014, according to the ICO's investigation, Cambridge Analytica wanted to take advantage of the pre-existing access to Facebook data enjoyed by the application developers with access to Version-1 of Facebook's API as their primary objective involved a planned use of this data in order to create data models which would inform on their work on electoral campaigns in the USA. Since Cambridge Analytica did not have access to the Version-1 at the time because of no pre-existing application on the platform, therefore in order to gain access Dr. Kogan arranged and proposed a merger between Dr. Stillwell's team in order to serve their personal interest of accessing facebook user's personal information.

> "CA initially discussed a collaboration with Dr. Stillwell, since Dr.Stillwell's application 'My Personality' had already collected a huge amount of Facebook data legitimately for academic purposes. Dr. Stillwell refused CA's offer, citing data protection concerns."" (ICO, July 2018)

Various employees working for Cambridge Analytica came forward and served as a catalyst during the whole investigation. On numerous accounts, witnesses have told the ICO about the brilliant strategy Dr Kogan employed to gain access to Facebook friend data on version 1 of their graph API.

> "Witnesses have told us that in order to gain access to Facebook friend data on API V1, CA initially discussed a collaboration with Dr David Stillwell. Dr Stillwell's app, 'MyPersonality' had already collected a large Facebook dataset – this data was legitimately collected for academic purposes. Dr Stillwell refused CA's offer, citing data protection concerns as his reason for not allowing the company access to the MyPersonality dataset." (ICO, July 2018)

During the May of 2014, Dr. Kogan, with links to Cambridge University, offered to undertake the work himself as he had already developed his own app called the 'CPW Lab App' later renamed as 'Thisisyourdigitallife' which was operating on API V1. Further evidence suggests that CA staff assisted Dr. Kogan to set up Global Science Research, once being set up a contract was signed with Cambridge Analytica. Information reviewed by the ICO implies that in order for a Facebook's user data to be harvested and processed by CA, the user or one of their Facebook friend, would have had to log into and authorize the app, The data of these users and their Facebook friends was then available to GSR and ultimately to Cambridge Analytica."app accessed up to approximately 320,000 Facebook users to take a detailed personality test while logged into their Facebook account.

> "However, in 2014, Dr. Kogan was introduced, via a colleague who knew Mr. Wyile, to SCL Elections Ltd, which it is believed was interested in the "My Personality" application. Dr Kogan approached others at the Psychometeric Centre about the possibility of a commercial venture with SCL Elections Ltd but the other academics at the centre decided not to participate on the terms involved" (ICO, July 2018)

The witnesses involved Cambridge Analytica staff members, including the whistle blower Chris Whylie, as they were involved in setting up these contacts through their networks of friends and colleagues; many of whom were involved in earlier political campaigns in North America.

The sections in (Figure 12) represent a typical Facebook user's personal information which unique to each and every active individual user. Defined clearly in the EU's General data protection regulation as,

- Personal data is information that relates to an identified or identifiable individual.
- What identifies an individual could be as simple as a name or a number or could include other identifiers such as an IP address or a cookie identifier, or other factors.
- Even if an individual is identified or identifiable, directly or indirectly, from the data you are processing, it is not personal data unless it 'relates to' the individual.
- When considering whether information 'relates to' an individual, you need to take into account a range of factors, including the content of the information, the purpose or purposes for which you are processing it and the likely impact or effect of that processing on the individual. (ICO, May 2019)

A.2 Accessing Facebook's friends data.

Dr Kogan's online personality assessment application had access to a user's **Posts, Photos and Videos, Comments, Likes and Reactions, Private and Group Messages, Profile Information** along with this, **their name, gender, age, current city, new feed posts, friend list and email addresses** were collected. In addition to the data collected directly from the personality test itself. The app utilized the Facebook login in order to request permission from the app user to access certain data from their Facebook accounts. As a result, the app was able to collect the following categories of information from the user to varying degrees. (ICO, July 2018) Cambridge Analytica commissioned another third party survey company called Qualtrics who then sought out and paid the members of the public, less than a dollar to access the application, this step was taken in order to maximize the number of Facebook User's data which was accessible to GSR and, ultimately to Cambridge Analytica. Once the data was obtained by Dr. Kogan's company GSR, it was then modelled and transferred to a secure 'drop-zone' from where Cambridge Analytica was able to extract the data relating to the data subjects that were interested in and for whom they had pre-existing data. Further data modelling was performed by the data scientists over at CA in order to create 'proprietary data models' which was then used during their political targeting work in the US.

> "In the course of these actions ICO seized significant volumes of evidence, including mobile telephones, storage devices, tablets, laptops, numerous servers, financial records and paperwork of relevance to our inquiries. At one location we discovered a number of disconnected and physically damaged servers; these servers have been subject to intense digital analysis to recover relevant material at component level." (ICO, July 2018)

A.3 Factual evidence: Exchange of Data between GSR and Cambridge Analytica

This whole investigation opened up a pandora's box and raised once again a chilling concern regarding the very definition of 'privacy' in the digital age. This investigation, in one way or another serves a mere reflection on the ethical and moral conduct of data controllers and processors, at an interpersonal and as well as at an organizational level. The information commissioner's office accumulated evidence suggesting that the Cambridge Analytica staff helped Dr Kogan set up Global Science Research Ltd. Upon the the set up, once it was finalized and functional, Dr Kogan signed a contract with Cambridge Analytica.

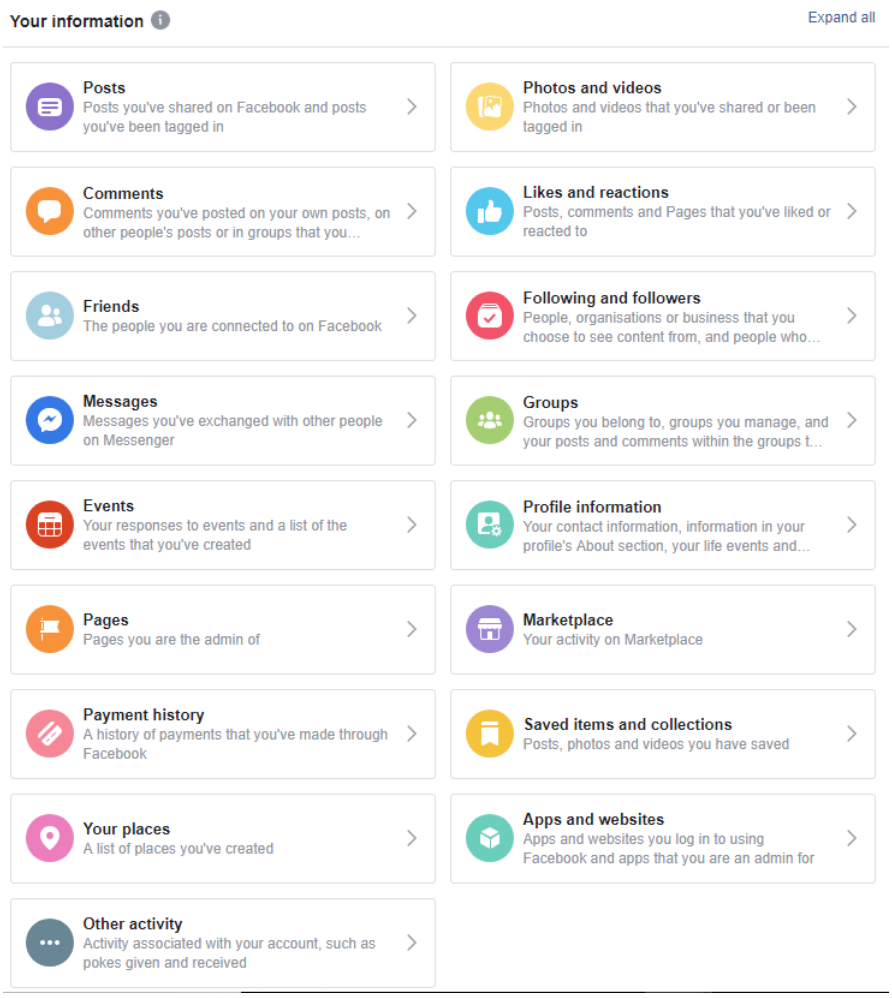

Figure 13: Facebook user's personal information

> "The ICO has evidence that CA staff assisted Dr Kogan to set up GSR. Once the company was set up and a contract signed with CA, Dr Kogan, with some help from Chris Wylie, overhauled the 'CPW Lab App' changing the name, terms and conditions of the app into the 'GSR App' which ultimately became thisisyourdigitallife (the app). Information reviewed by the ICO suggests that in order for a Facebook user's data to be harvested and processed by CA, the user, or one of their Facebook friends, would have had to log into and authorise the app. The data of these users and their Facebook friends was then available to GSR and, ultimately to Cambridge Analytica"

The application's name was changed three times before it was finalized. The contract between GSR and Cambridge Analytica served as a bridge to transfer user personal data from Facebook, via GSR and finally to Cambridge Analytica.

During one of his scarce appearances, before the DCMS Select Committee, Dr Kogan explained how GSR built data models predicting how a user is most likely to vote. In his testimony he told the Committee that GSR took a Facebook user's answers to the application survey which were used to make a certain predictions about the user; combined with other information taken from the user's profile (Figure-4) gave GSR detailed and accurate insight regarding the voting behaviour of population that they were dealing with. According to the ICO's report into the investigation published in July 2018 they stated,

> "because of the configuration of API V1, GSR also received the public profile information about the app users' Facebook friends, including their Facebook likes. As such GSR was able to provide modelled data about the 'app' user and their Facebook friends whose privacy settings allowed access by third party apps." (ICO, July 2018)

During the early stages of inquiry, Facebook was alerted regarding the breach; by the ICO but initially by media coverage during the mid 2015. The ICO enforcing the powers under the UK data protection

act-1998 managed to obtain a warrant for access to the premises of Cambridge Analytica. During the extensive 7 hour search, ICO recovered a substantial amount of evidence that pointed towards the legitimacy of the illegal data transfer between GSR and Cambridge Analytica.
The seized evidence suggest an approximate of 700 terabytes of personal data have been transferred from GSR to Cambridge Analytica. The ICO, as a part of the investigation inquiry with the staff at Cambridge Analytica confirmed that even though a little effort was made to delete the data acquired by GSR when Cambridge Analytica was first contacted by Facebook. It surprisingly took the social media giant an investigation by the ICO to wake up and establish a dialogue with Cambridge Analytica. However, unfortunately according to an updated version of the initial investigation report states,

> "some 'proprietary data models', data models derived from the data harvested from Facebook, may not have been deleted. We will be making sure any organizations, which may still have copies of the Facebook data and its derivatives demonstrate its deletion." (ICO, Nov 2018)

Facebook, during a hearing of the Department for Digital, Culture, Media and Sport select committee provided clarification in response to the chairman Damian Collins objections regarding a complaint filed by the US Securities and Exchange Commission (SEC) which indicated that Facebook staff knew about the data being compromised earlier than its senior staff acknowledged to the MPs. This raised concerned about Facebook providing contradictory evidence to the parliament. The complaint claimed that,

> "While Facebook initially raised concerns on the matter in September 2015, Facebook executives said the site first learned of the data misuse months later.[64]" (The Guardian, 12 Aug 2019)

In their defense, the Facebook's UK head of public policy, Rebecca Stimson said, "the company and its staff truthfully answered the questions and rather than contradictions, provided accounts of two separate events."

[64]Kevin Rawlinson, 12 Aug 2019, Facebook denies giving contradictory evidence to parliament, The Guardian, Available at: [https://www.theguardian.com/technology/2019/aug/12/facebook-denies-giving-contradictory-evidence-to-parliament]