PART I

Introduction: Theoretical Perspectives

1

The Pixelated Person: Humanity in the Grip of Algorithmic Personalisation

Uta Kohl

1.1 INTRODUCTION

By far the most fascinating and profitable subject of predictive algorithms is the human actor. The capacity to predict human preferences, responses and behaviours offers endless possibilities for science, commerce, politics and regulation, and promises convenience and efficiency that further private and public interests in equal measure. There is nothing inherently new about the attempt to predict buying choices, political leanings and likely votes of individuals and groups, the probable effectiveness of medical treatments, likely defaults on loans, the chance of fraudulent insurance claims or of reoffending. Yet, the capacity to 'know' the individual and the group, and to predict their constitution and behaviour has witnessed a sudden upturn of unprecedented scale. The rise of network society and smart technology is generating endless trails of personal data, finely pixelated digital footprints, that are aggregated into big data sets – that involve large collections (volume) of real-time (velocity), diverse and relational personal data (variety)[1] – about virtually all aspects of human life from shopping, food and entertainment preferences, friendship networks, romantic attachments, social activities, health concerns, physical movements, driving behaviour or sporting activities, to biometric data, such as voice, face, gait or keystroke, or physiological data on heart rate, blood pressure or sleeping patterns. These data sets, when mined by algorithms, can reveal significant patterns and correlations and, ultimately, produce knowledge about the group (e.g. behavioural trends, economic activity, delinquency, spread of disease, political trends, etc.[2]) *and* about the individual (e.g. educational level, social status, political leaning, sexual orientation, emotional states and psychological vulnerabilities as well as predilections for activities and movements). This knowledge then lies at the disposal of the private sector and government to be used for a wide range of purposes, implemented through 'personalised' services, treatments and regulation – some beneficial, some harmful, but mostly a mixture of both.

For example, only a few Facebook 'likes' are needed to reveal correlations with personal attributes, such as sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age and gender.[3] These granular insights into the individual or micro-groups can be, and are, used to select and deselect content and advertisement to match their profile. Such 'personalisation' serves as essential information management (in response to the overwhelming amount of information available), and promotes efficiency (by saving users' time in searching through masses of content, and businesses the expense of serving adverts to uninterested users). For many, personalisation offers the customisation and optimisation previously only available to the elite, e.g. the personal advisor or trainer, whilst the great masses had to be content with mass production. Mass-personalisation or individualised consumption at scale is now possible at least in the service industry.[4] Yet, the very same practices that appear so beneficial show their exploitative dimension when used to extract extra value from the consumer as, for example, when an inferred desperate search for a loan is translated into an offering of credit with a 'personalised' higher interest rate that reflects the urgency of the search. Equally, the manipulative aspects of personalisation shine through in the practice of micro-targeting political adverts to profiled users, and undecided voters, in the lead up to elections or referendums, as revealed in the Cambridge Analytica scandal.[5] Although the scandal centred on the deceptive collection of data and the absence of user consent to such collection and use, consent seems to only marginally address the manipulative inflection of political (and other) micro-targeting. Even where targeting is consensual, the 'opted-in' lack of choice and consequential lack of exposure to alternative narratives still seem problematic. By the same token, if a patient's personal medical history is supplemented by a genetic profile from an ancestry service, like 23andMe, and life-style data from a Fitbit watch in order to decide on the most effective made-to-measure medical treatment,[6] this process seems in the patient's interest (most effective treatment) and in the public interest (efficient allocation of scarce resources). Yet, the same practice becomes more suspect when used to limit otherwise available treatment options or deny treatment altogether, on the basis of unfavourable DNA or life-style profiles. Finally, the possibility of predictive policing through micro-segmentation of populations promises to employ scarce police resources more efficiently by concentrating on likely serious delinquents and thus to pre-empt crime and disorder more effectively. Yet, he who seeks finds: the distorting impacts of such targeted practices have been well documented, and one of their concomitant side effects is that some sections of the population are granted leeway from which others do not benefit, often along historic racial and ethnic lines of division.[7]

Whether beneficial or detrimental, what these scenarios have in common is the data-driven profiling of consumers or citizens to deliver a customised or personalised service, advert or legal response. Personal data in conjunction with big data is interpreted by algorithms to create a picture of who someone *is* based on who they

were – their past preferences, activities, networks and behaviours – in order to make a future-oriented prediction of what they might like (i.e. which film), what might persuade them (i.e. which ad) and how they might act (i.e. commit a crime or succeed in a job). A key problematic of profiling and customisation practices lies in their very virtue: the pre-selection and pre-emption of individual choices by those with access to big data sets and profiling technology. Thaler and Sunstein have called them 'choice architects' in the context of 'nudging'.[8] The pre-management of individual choices by these architects is rendered at times more benign by the triviality of the personalised service, e.g. a recommended book, film or song; or by the perspective of those upon whom personalisation bestows a benefit based on their 'good' profiles, e.g., the healthy patient, the unlikely delinquent, the creditworthy or price-sensitive consumer.[9] The core of the problem, however, remains the same and lies, first, in taking the human agent out of the loop of participating and directing her individual and collective life through making active choices,[10] in potentially two capacities: one, the algorithm replaces the traditional human decision-maker (e.g. the judge or the editor or the business person) and, two, those decisions then also pre-empt the choices of the profilee (e.g. the defendant or consumer). Choices are made for her, or at least the framework is created within which she can make her choices. The *second* problematic underlying personalisation practices is that big data analytics is generated by autonomous technology whose complex processes, optimised through feedback loops and machine learning capabilities, often place it beyond human comprehension, and thus *prima* facie also outside human oversight and contestation.

This collection of essays engages with these problematics in various social domains and academic fields of inquiry, and brings together scholars from different walks of law (data protection and privacy law, criminal, medical, and contract law as well as constitutional theory) and other social sciences, such as political theory, human geography, criminology, behavioural economics and philosophy, to interrogate this new powerful phenomenon that is sweeping across economic, political, social and legal domains, and dramatically reconfigures our social structures. What is striking about the contributions is that, despite the different contexts and perspectives, persistent themes emerge. On a practice-focused level, data-driven profiling and its myriad uses raise questions about *substance* (e.g. what is the accuracy of the profile and the legitimacy of using probabilistic predictions in favour of, or against, an individual, particularly in light of the possibility of mistakes or discrimination; what are the wider unintended consequences of profiling on private and public or collective interests) and about *process* (e.g. what oversight, if any, is exercised over the autonomous decision-making technology?; can informed consent ensure the empowerment of users in their profile creation and, more generally, to what extent can and should individuals be able to resist and challenge the collection of their data, its aggregation and use?; and how does it impact on avenues for collective resistance?).

On a theoretical level, there are also clusters of ideas that cut across subject-matters and disciplines, and flock around two themes. The first focuses on the foundational premise of predictive technology which is that future actions can (and should) be inferred from past behaviour or from the behaviour of like actors – a premise which is at odds with ideas of moral agency and free will. Yet, agency lies at the heart of our social orders and underpins the *homo economicus*, the self-determining citizen, and the moral actor who can only be held responsible for their actions on the basis of the freedom to act otherwise. Moral agency is also closely related to our conceptions of identity and personhood, and the open-ended evolving nature of human individuality. These conceptions are profoundly challenged by the creation of the *pixelated human* – a digitally constructed, two-dimensional, instrumentalised, commodified representation of individuality – and yet, this entity is frequently treated as the authentic self. Furthermore, under this deterministic view of human behaviour, normative questions are reduced to, or disguised behind, empirical observations about individual and group histories. This essentialist approach has the effect of continuously reasserting the status quo, and thereby consolidating and exacerbating it, including existing inequalities, structural disadvantages or political world views, and concomitantly reducing the room for individual or collective betterment.

The second cluster of ideas places the granularly profiled user from whom value can be extracted (generally in the name of efficiency) within a sharpened capitalist economic order. Shoshana Zuboff argued that the new data practices have given rise to surveillance capitalism: 'surveillance capitalists discovered that the most predictive behavioral data come from intervening in the state of play in order to nudge, coax, tune and herd behavior toward profitable outcomes.'[11] This perspective helps to frame the heightened user-pay model that various personalisation practices (e.g. personalised health care, credit or insurance products) implement as instantiations of liberal ideas of individualist fairness or just desert in opposition to notions of communal solidarity or distributive justice.[12] The free market lens also helps to explain why consent and personal autonomy should so systematically underwrite profiling practices, regardless of the facts that users exercise that autonomy within vastly asymmetrical power relations; that it legitimises value extraction as opposed to offering protection; and that invariably more is at stake than individual private interests. Equally the commodification of personal data is only intelligible against market logic. When consumers can sell their personal data in return for 'free' services, and corporations can buy and ring-fence this vast resource, the potential of these data sets as a (global) public good to be used for the benefit of all becomes much more circumscribed.[13] At the same time, the micro-segmentation of communities through personalisation practices, legitimised by individual consent, fragments political communities and distorts democratic processes, with the compounding effect of weakening a key mechanism for holding corporate and governmental actors to account, and for restraining the very processes that undermine those democratic processes.[14] In short, profiling and personalisation

practices are deeply inscribed with capitalist market values – from their initial conception and rationalisation to their implementation within economic, social and political spheres and their continuing legitimation.

If there is one theme that carries through the whole collection, it is that this newly emerging and highly disruptive phenomenon has continuities with previous practices, concepts and ideologies, through which it may be analysed and critiqued. It is also only against these previously established understandings and processes that we may recognise how it presents a paradigmatic shift that really deserves our assiduous attention before it has pervasively and conclusively reshaped our social orders in its own image. This introductory chapter provides reflections on two distinct intellectual hinterlands to the more specific themes and applications of data-driven personalisation practices in this collection. First, it situates these discussions against a general framework of profiling and defends data-driven individual and group profiling against some critiques of stereotyping, on the basis that our cognition of the external environment is necessarily reliant on relevant abstractions or non-universal generalisations. The second set of reflections centres around the philosophical tradition of empiricism as a basis of knowledge or truth production, and uses this tradition to critique data-driven profiling and personalisation practices in its numerous manifestations. The final part of the chapter summarises the chapters in this volume and their individual contribution to the overall narrative.

1.2 INDIVIDUAL AND GROUP PROFILING AND THE VIRTUES OF STEREOTYPING

1.2.1 *The Interdependence of Individual and Group Profiling*

An initial controversy surrounding algorithmic profiling based on large sets of digital footprints is whether the individual or the group is its real target and the potential object of manipulative practices. Whilst the language of personalisation and customisation suggests the individual is the focal point, in some ways 'personalisation' is a misnomer, as individual profiling is always a form of classification whereby the individual is assessed against group attributes (more on that below) and then put in a micro-category for the purpose of delivering the 'personalised' response or service. Thus although the individual is the target of the customised message, service or treatment, the outcome is based on group features and multiplied across the micro-group. Furthermore, the fact that individual profiling is premised on analysing data sets about populations – mined for correlations and leading to the construction of groups in the process – has led some to conclude that *group* profiling is the critical new phenomenon that challenges existing legal modalities:

> The search for group privacy can be explained in part by the fact that with big data analyses, the particular and the individual is no longer central. In these types of

> processes, data is no longer gathered about one specific individual or a small group of people, but rather about large and undefined groups. Data is analysed on the basis of patterns and group profiles; the result is often used for general policies and applied on a large scale.[15]

This argument has some validity (given that privacy regimes envisage an individual victim, and harm to the group only derivatively), but the ability to micro-profile individuals is still at least as valuable to corporate and governmental actors as knowledge about the group, as borne out by the widespread emergence of personalisation practices. In any event, the individual-versus-group dichotomy may largely be misconceived because they reflexively interact with each other. Individual data feeds into population data sets and these sets produce, through correlations, knowledge about populations, that is patterns and groups within them (inductive), which in turn are instructive about the individual (deductive).

The close, yet varying, integration of individual and group profiles has been subject to some debate and conceptualised in the distinction between distributive and non-distributive group profiling.[16] For *distributive* profiles (universal generalisations) attributes of the group are 'actually and unconditionally manifested by all the members of that group'[17] and thus group membership also allows for definitive inferences about the attributes of its members.[18] Every member of university staff (the group) has an employment contract with the university and a salary (attributes). In contrast, *non-distributive* profiles (non-universal generalisations or stereotyping) refer to groups where a family resemblance unites members, but not every member shares every attribute.[19] Here 'a group is defined in terms of… *significant deviances from other groups*. They are based on comparisons of members of the group with each other and/or on comparisons of one particular group with other groups.'[20] The group boundaries in non-distributed profiles are inevitably fuzzy. Those with a high risk of cardiovascular disease (group) share a number of risk factors, for example, lifestyle, genes, age, weight, etc. (attributes),[21] but membership does not allow for definitive inferences about the particular attribute of a particular member. The non-universal generalisation that 'young men drive recklessly' does not allow for a definitive inference about the driving of any particular young man but, as argued below, mistakes on the individual level are often legitimated by the benefits of identifying (empirically sound) tendential truths.

Whilst non-distributive profiling explicitly compares one group *vis-à-vis* other groups, ultimately the distinctiveness of a distributive group profile (university staff) can also only be understood against other groups, that is what it is not (police or hospital staff, or university students). Indeed, the difference between these two types of profiling may in practice (and theory) not be that clear cut (i.e. is the whiteness of swans 'necessarily manifest' or non-essential?) and becomes largely a function of the profiler's knowledge, pre-conceptions and attendant construction of the group. This suggests that the certainty of (empirically based) distributive profiles may be illusory.[22] The two types of profiling may simply reflect different philosophical traditions: distributive profiles adopt a Platonic top-down perspective on a concept or class that assumes and finds a common

essence underlying all its manifestations, whilst non-distributive profiling builds on Wittgenstein's bottom-up (and empiricist) notion of family resemblance whereby concepts or words just refer to clusters of similar or related phenomena:

> Consider for example the proceedings that we call 'games'. I mean board-games, card-games, ball-games, Olympic games and so on. What is common to them all? – Don't say: "There *must* be something common, or they would not be called 'games'" – but *look and see* whether there is anything common to all. – . . . [W]e see a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail..[23]

So arguably distributive and non-distributing group profiling does not refer to different types of groups, but rather to different ways of looking at the same group, or, more precisely, to different ways of *constructing* groups.

As non-distributive profiling can capture a wider range of relevant, albeit non-essential, attributes (as opposed to seeking a group's essence), it yields a much richer picture of groups and individuals, but also has blurry edges and is fallible in respect of making definitive inferences about its members.[24] This is significant for big-data individual profiling, or any form of statistical profiling: when individual profiles are inferred from comparison with the group (indirect profiling), it may be tempting to fill 'gaps' in an imperfect overlap with the missing group attributes. For example, in the policing context, a large aggregated criminal justice database with data on criminal activities mapped onto post codes, on criminal records and recidivism, social media activities and networks, education and employment histories of offenders, and personality traits may – based on strong correlations – predict for a particular offender a high risk of recidivism. The Harm Assessment Risk Tool, or the HART algorithm, used by Durham Constabulary makes such predictions based on 509 'votes' by the system.[25] A digital footprint on social media may, in the absence of explicit evidence, be analysed to 'reveal' the missing attribute of a single person's status, a left-wing political outlook or homosexuality. Based on the strength of the correlation, an unknown attribute may be 'highly likely' and in this respect fall somewhere between the distributive and non-distributive profiles – as neither necessarily manifest nor simply possible. However, the effect of the use of the predictive technology will often be such as to treat highly likely predictions as effectively established, along the lines of Plato's essentialism. Yet, there may be rights-based reasons, such as the presumption of innocence or the right to privacy, why a particular inferred attribute should be treated as non-essential and its absence presumed, as, for example, when sensitive data may be inferred from a range of non-sensitive data points. (Chapter 5)

1.2.2 *The Virtues of Stereotyping*

One persistent objection to individual profiling based on comparisons to group data, including big data profiling, is that the resultant stereotyping (or non-universal

generalisations) leads to the 'deindividualisation of the person' which occurs when '[p]ersons are judged and treated more and more as members of a group (i.e. the reference group that makes up the data or information subject) rather than as individuals with their own characteristics and merits.'[26] This critique is directed at *indirect profiling* that draws inferences about the individual from group data (invariably through non-universal generalisations), as opposed to *direct profiling* that is ostensibly based on data only about the particular individual and therefore arguably more accurate.[27] The objection to the 'deindividualisation of the person' or stereotyping based on comparisons with the group is flawed for a number of reasons. First, the argument that direct profiling delivers *prima facie* more legitimate profiles as it is solely focused on the digital footprint of the single individual assumes that someone's past activities and preferences provide a valid yardstick for his future behaviour and preferences, and implicitly assumes that personhood is fixed in time. Such reasoning relies as much on stereotyping of the individual (and on denying agency) as indirect profiling, as it does not allow for the possibility of continual reinvention and development of individuality through repeated assertions of free choice. Indirect profiling has at least the virtue of squarely acknowledging that 'no man is an island' and that individuality is intimately tied up with social forces within which it develops and against which it may be understood. Still, *all* profiling used for predictive purposes is inherently irreconcilable with the notion of free will as underwriting moral and legal responsibility as well as autonomous participation in democratic processes.

Second, direct profiling is also necessarily comparative with the group, much like indirect profiling, and cannot but invoke the social dimension of human existence. Individuality can only be understood against an assumed 'normality' which contextualises individual divergence.[28] An individual's social media digital footprint is entirely meaningless by itself, in a social vacuum. It can only signal depression or creditworthiness or criminogenic tendencies against data sets, drawn from the group, that display the whole spectrums of psychological, financial or criminogenic states.[29] The interdependence of the individual and the group, the particular and the general, uniqueness and commonality, may best be illustrated with reference to DNA profiling as the biological equivalent to behavioural profiling:

> DNA fingerprinting (also called DNA profiling or forensic genetics) is a technique employed by forensic scientists to assist in the identification of individuals or samples by their respective DNA profiles. Although more than 99.1 per cent of the genome is the same throughout the human population, the remaining 0.9 per cent of human DNA shows variations between individuals.[30]

In parallel with the biological profile, where commonality far outweighs uniqueness, and individual genomic variations operate, and are identifiable, against genomic commonality, individual behavioural uniqueness can also *only* be conceptualised against the broad brush of collective humanity. The specific and the general are co-dependent. (See Chapter 5.)