Institutional Exclusion after the Introduction of Big Data Systems in India

Leiden University

Supervisor: Prof. dr. N. K. Wickramasinghe

Bachelor of Arts; International Studies

Thesis Tutorial Group: South & Southeast Asia

By: Luke Caradine

Student Number: S1231871

Word count: 10894

Date: 28-05-2015

Abstract

The means of identification of the population determine the institutional practice. The census in India led to the institution of the caste-system in order for the British to better manage the population. Under the Foucauldian governmentality paradigm, this thesis argues that the modern practices used by the government for identifying the population are a continuation of the earlier census method of identification. The extrapolation of characteristics to determine the structure of Indian society was a system that relayed information from the population back to the government through the census. Segregation or exclusion of people who fell outside of the proposed structure presented in the census were homogenised within castes. Contemporary systems for the population to relay information to the government may move beyond the earlier traits of structure. However, as the government becomes increasingly dependent on technology to identify and gain insight into the issues facing the population, the role of technology in providing the government with data becomes an issue. Essentially, the government is in a process of increasing automation in institutions in order to better assess the population. It is capable of putting in place the systems that register the issues facing the population. Yet, as the institutions begin to provide an assessment of the data to the government in order for them to act. Those people within the population who are not registered in- or providing relevant data to – institutions therefore fall outside of the assessment of the government. As they do remain part of the population this thesis will argue that the big data systems (1) are dependent on the means that the government to collects data, (2) do not include the whole population and thus policy recommendations, based on big data assessment, require the government to extrapolate the perceived issues to the whole of the population, (3) the welfare of the population as the end-goal of the government will see a changing role in citizenship when the policy of the government becomes increasingly determined by the information the citizen provides.

Contents

Introduction

The Case Study

The largest biometric identification system in the world is part of the collective institution Unique Identification Authority of India (UIDAI). In 2009, India, in an attempt to standardize the identification system, put in place a body of agencies that collectively issue an 'Aadhaar' card to carry out the Unique Identification (UID) program. Its goal is to provide every Indian citizen with a unique 12-digit card containing a microchip wherein is recorded, at least, the individual's fingerprint, iris-scan and existing demographic information. The vast processing power available in the digital world can carry out a task never before attempted on this scale. In the UID program the massive empty data space in large server towers provides the government with a means of recording information on citizens. Essentially, the UID program will be able to provide a number per citizen rather than a host of diverse cultural, political, status-related or religious labels. In the last few decades a multitude of scholars, politicians and activists have been attempting to assert the need for the government to recognise and deal with areas and groups differently. The government in turn has put in place an institution that attempts to reconcile the heterogeneous society with a means of identification that surpasses the heterogeneity and homogenises numerically. The reconciliation between the heterogeneous society and the current political machinations is what is coming to the fore now. The Aadhaar initiative in its own words is "devoid of caste, creed, religion and geography". By establishing a system that gives the individuals within the population a number rather than say number with a series of letters which indicate caste or religion the Indian government is attempting to overcome what differentiates citizens by unifying them within a single system.

The largest biometric identification system in the world provides the registered citizens with an identification card that houses their unique characteristics. As a system that creates an overarching method of identification, rather than a system that attempts to incorporate the

various existing local agencies within a larger whole, it is able to immediately establish a top-down approach for identifying citizens. Such a system necessitates the creation of local agencies that are handed and taught the method for incorporating the citizens whom are registering with the UIDAI, the Enrolment Agencies. The local enrolment agencies serve to reduce or supplant the existing institutions and agencies that provide residency identification. It is a means of the government for solving the "problems of inefficiency, corruption and fraud endemic in the existing system, in which overlapping jurisdictions resulted in up to twenty different forms of identification issued by various local and national agencies" (Schmidt and Cohen 78). Current estimates indicate that 67% of the population[1] has been registered by the UIDAI and been issued an Aadhaar card (The Economic Times). This means that approximately 817.8 million people have had their biometrics recorded and given demographic information to this institution. Although the card is not compulsory the UIDAI has made some connections to other institutions, such as educational institutions, that are making the card a vital tool for people to access public services. Taking two of these potential connections from the main UIDAI website on education and public health there is an indication of current importance and potential future developments of the Aadhaar card. In education the document reports that: "Provision of UIDs will ensure that there are no problems due to migration of students anywhere within the … [it] will effectively address the issue of education of children of migrant labor as their children can be admitted at new place without any other verification" (Unique Identification Authority of India). In the public health document the UIDAI claims that: "by linking citizen ids with hospital or other medical facility records generated through facility visits can (1) inform the public health system of the prevalence of various routine disease conditions (2) help prepare the health system to respond to unforeseen epidemics" (Unique Identification Authority of India). The progression of the systematic incorporation of

[1] A figure based on the 2011 census

citizens is proceeding alongside the attachments the UIDAI is making or proposing to other institutions. Therefore, it is likely that the Aadhaar card will come to take on additional roles next to simply being a means of identification. In particular, having the Aadhaar card for an internal migrant worker and his or her family may come to determine the educational support offered to children and, potentially in the future, the healthcare received.

In a speech given to parliament in 2012 the president of India made clear the government's perspective on the role for the Aadhaar card. He stated in point 8 that:

"*To reach the millions of underprivileged people, my Government has launched a unique ADHAAR*[2] *scheme which would help improve service delivery, accountability and transparency in social sector programmes and lead to their financial inclusion*" (Press Information Bureau, Government of India).

Point 8 above is a continuation of point 7 which highlights the government's decisions to reduce corruption. Here the president stated that:

"*Efficient and automated delivery of public services with minimum human intervention is one of the keys to reducing corruption. Under the National e-Governance programme, more than 97,000 Common Service Centres have been established across the country for making public services conveniently available to citizens*" (Press Information Bureau, Government of India).

The role of Aadhaar illustrates the government's idea of increasing automated delivery of public services. The Aadhaar card and information stored therein has the potential to automate response from particular institutions and organisations. However, there is an implicit

[2] I presume the press bureau wrote it wrong.

consequence herein for the people who choose not the register. This is demonstrated by the National Electorol Rolls Purification Authentication and Programme (NERPAP). NERPAP is an example of an institution that is attempting to reduce corruption. In particular, NERPAP is attempting to reduce voter fraud or multiple votes by a single individual[3]. One of the key criteria for being registered in this program is the Aadhaar card (National Electrol Rolls Purification Authentication and Programme). This indicates that the Aadhaar card is becoming increasingly important not just for the government's ability to "improve service delivery, accountability and transparency in social sector programmes and lead to their financial inclusion" but is also becoming a necessary feature for the citizen to access these services.

Thesis Statement & Lay-out

The role of the systems that provide the automated service delivery to citizens is becoming a feature of many countries and organisations. Although in India the Aadhaar card and the UIDAI set out to reduce inefficiency, corruption, cultural distinctions and other goals, the use of these massive initiatives is marked by a contemporary trend in digital opportunity. The trend is one more commonly referred to as Big Data and its usage. It is becoming more prominent in society and governments as "big data refers to things one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value, in ways that change markets, organizations, the relationship between citizens and governments, and more" (Mayer-Schönberger and Cukier 6). The notion that big data can 'extract new insights' to 'change the relationship between citizens and governments' takes precedence in this thesis. In particular, as E. Schmidt and J. Cohen argue, "for governments […], this thriving data set is a gift, enabling them to better respond to citizen[s]" (Schmidt and Cohen 57). The Aadhaar card is indeed an example of such a new big data method of governments to change

[3] Whereby, for instance, the voter moves from on constituency to another to vote there with the same name. As per the Representation of the People Act of 1950, India has made it illegal to vote in more than one constituency.

the relationship between governments and citizens. Recording data on near 816.7 million people may show the government certain patterns in the population. More so as new information is tied to the Aadhaar card when new institutions come to collaborate with the UIDAI (as the NERPAP has done). Alternatively, the case made for linking the Aadhaar card to public health institutions to oversee the spread of disease and monitor epidemics indicates how the big data systems may come to monitor future trends in the population. However, the 67% of the population that has an Aadhaar card or has been registered by the UIDAI does not cover all of the population of India, nearly 400 million people are not in this system. Although the use of these big data systems have the potential to indicate beneficial patterns for institutions there is still a large portion of the population who are excluded from these patterns. Big data systems used by institutions have at the core the notion that technological development and more encapsulating information gathering tools will have an effect on the citizens. Although how this will impact citizenry when there is a group outside of-, or excluded from-, the current big data system is the topic of interest here. It is in this debate that this thesis will attempt to contribute an argument. It will investigate to what extent big data systems used by governmental institutions allow exclusion; or conversely to what extent big data systems include marginalised populations in support of institutional policies. I will first develop a historical account of how identification practices affect the political structures in India. Second, I will look at how the census as a tool to identify the population can be understood under Foucault's governmentality paradigm. Third, I will link the governmentality paradigm to the big data systems and the opportunities that these provide institutions. Finally, I will argue that under the governmentality paradigm the information being relayed between the population and the government is increasing allowing the government to better develop policies. However, I will also argue that the institutions who are integrating big data systems are becoming reliant on the scope of the data supplied. Thus, there may be a continued rationality of exclusion.

Historical Development of Identification Tools

Linking the Modern to the Historical

The UIDAI has set up a big data system for recording all of India's citizens in one single database. It is attempting to move beyond a heterogeneous society by issuing the 12-digit number per citizen for all citizens. The labels that may have been previously attached to a group and by extension the individuals, have been removed and replaced by numbers that carry no, for instance, cultural identifiers. By removing these identifiers the government is put in a position whereby it is able to create segments of the population from the whole that are homogenised by a particular category which they themselves elect to operate on. Thus, the government is able to segment the population into, for instance, categories of mobility, parentage or age. Governing the population, then, becomes a matter increasingly linked to the technological support and techniques for collecting and assessing information that is available to the government. The government is able to use the Aadhaar system as a means to develop social service plans. As E. Jacobsen argues: "In India, biometric identification is to provide the foundation for a wide array of governmental schemes, including cash transfers and welfare distribution programmes, as well as surveillance" (Jacobsen 457). Moreover, the system is moving beyond the economic incentive. In an attempt to gain a better recording of the homeless or mobile population in New Delhi a United Nations Development Programme sponsored a large scale survey conducted by the Department of Social Welfare in 2009-10. Simultaneously, there was a registration with the UIDAI and Aadhaar cards were issued. Jacobsen writes that "the implementation of biometric identification technologies is thus an attempt to establish a more stable overview of these mobile targets of governance" (Jacobsen 465). The biometric system here reflects on how the government is using technology to create new more encompassing means of identifying the population to assess the means of governing.

The Census

The development of the caste-system reflects how the British government chose to develop its means of governance over a population based on a statistics. Dirks and Said argue, the creation of the caste became an exercise of using empirical information. Dirks writes that "it was the decennial census that played the most important institutional role not only in providing the 'facts' but in installing caste as the fundamental unit of India's social structure" (Dirks 68). While Said argues a similar point by way of asserting that the structures created by colonial powers were a consequence of the secularisation and the resulting imposition of "reconstituted, redeployed, redistributed" (Said 121) existential paradigms. It was, according to Said, "reducing vast numbers of objects to a smaller number of orderable and describable types" (Said 119) that lead to a classification of society. The early postcolonial thought has become a cornerstone for much of the discussion on the progression of the caste system. However, as N. Wickramasinghe points out there is a "need for a more contextualized, nuanced, and historically attentive approach to relations of power in colonial situations" (Wickramasinghe 34). It is in this line of thought that more recent scholars such as: Susan Bayly and Prachi Deshpande, have argued that there are certain issues that Dirks and Said do not take into account. Bayly argued that "in contrast to Dirks [there were] continuities between pre-colonial and colonial social orders" (Ballantyne 193); while Deshpande argues that "it was debates and struggles conducted in Marathi over the intersections between caste and regional identity rather than an all-powerful colonial state that […] shaped the politics of caste in western India" (Ballantyne 193). Regardless of the potential continuities that existed in the formation of caste it was fairly suggested by all authors that caste was a classification that in some way reduced society to homogenous entities.

Categorisation of Society

This is where the Said and Foucault position their notion of classification; "to know what properly appertains to one individual is to have before one the classification – or the possibility of classifying – all others" (Foucault, The Order of Things 157-158). Foucault develops this idea based on the development of natural history. He argues that the natural history tried to assert the 'autonomy of nature' and was thus in a conflict with the theological study of nature. The result was a break from the Classical rationality in theological logic with the intent to 'purify' and the "creation of a history that could at last be 'true'" (Foucault, The Order of Things 144). Moreover, he is, like Said after him, also arguing for the persistence of the rationality that existed during the Classical age into the rationality of the subsequent means of allocating order. Hence, when Said claimed that the colonial structures contained elements of a lasting existential paradigm he was referring to the notion of 'Exchange'. Foucault describes this as the synthesis of value and exchange "there would be no exchange if there were no immediate values – that is, if there did not exist in things 'an attribute which is accidental to them and which is dependent solely upon man's needs, as an effect is dependent upon its cause'" (Foucault, The Order of Things 215). Using the argument made earlier by Said the means designated necessary for a government to formulate how to govern a population are allocations of values that are 'accidental', in that they are allocated based on earlier or other attributes present elsewhere, and 'dependent' on man's (or a governments) needs. The persistence of earlier classifications that served earlier in Western countries can thus be argued to have been exchanged for a different articulation in India by creating a caste system but are inherently devoid of novelty. It is here that the caste-system as an institution becomes a perpetuation, albeit perhaps a continuation or a 'sped-up' version[4], of earlier Western rationality. Therefore, the chosen methods of governance, by instituting structural classification of society, are similarly based on the remnants of this same rationality.

[4] As Susan Bayly argues (Ballantyne 193)

The secular rationality that made up the census based approach for governing marks a step in the process toward the more contemporary development of big data systems. Although the census of the 19th century in comparison to features available today is not a particularly advanced technological feature, it does mark a means of collecting and assessing data on large portions of the population and acting on it. The census held a fundamental place according to Dirks in establishing a social structure that could be used by the government. He concludes with the argument that "under colonialism caste became a specifically Indian form of civil society, the most critical site for the textualization of social identity […] and the development of the documentation and certification regimes of the bureaucratic state" (Dirks 76). As such, the role of the census was used by a bureaucratic system to systematically differentiate between large groups of people. B. Cohn adds that "the 'official' view of caste was very much related to how the British collected information about the caste system […] the census, the constant need for government applications to identify oneself by caste, the application of varying law to different castes, all seemed to have played a part" (Cohn 15, 18) in the social change that was happening in India. N. Wickramasinghe summarises a continuation of one of Cohn's central arguments when she writes that: "the census, the survey and more generally the collection of knowledge that defines a 'population' […] can be used to maintain surveillance and superintend social change" (Wickramasinghe 36). The definition of the population, those characteristics that made up the caste, were used by the government to manage and maintain control. The official view of caste, established by census, was an integral feature of the British Empire. Robinson and Gallagher argue that the concept of the "official mind" is marked by how "in a strikingly homogenous political system, British statesmen and civil servants had worked together to assess available information and develop responses to policy challenges, guided by 'the long-run national interest which stayed much the same from ministry to

ministry'" (Potter 23-24). The structure imposed on society was thus an integral feature of the British Empire for at least ensuring the economic prowess of the empire.

Sharing Census Knowledge Globally

The economic aspect of colonialism is important in the maintenance of the system put in place in India. It is an argument stressed by C. Bayly[5], Ferguson[6] and Robinson & Gallagher[7] suggests that focussing purely on the social change dismisses the economic incentive in establishing empirical means of governance. The need to maintain a stable structure on society, one whereby the information and knowledge on that society could be easily passed on from official to official, suggests that there was a global process occurring. The focus on developing the tools to gather empirical data is an early indication of globalisation. The global reach of the British Empire and the power relations that were put in place in India suggest that there is a foundation for the necessity of the historical role empiricism as a means of gathering information. As the measurements of revenue in an Empire were often tied to a principal value, such as gold, so too the societal measurements and subsequent classifications were tied to categories of encompassing values. L. Dumont argues that the values that led to the creation of the caste institution in India are separated by time and contemporary ideals from the European countries, whom used similar values in the establishment of institutions. The role of the empirical means of assessing a society and subsequently establishing a system that manages and maintains it, is a feature that enables the maintenance of wealth. It is therefore in the interest

[5]C. Bayly argues that empire-building in South Asia was a means to gain financial support for military purposes (Potter 35)

[6] Ferguson argues that Britain enforced free trade in order for it to gain the financial supremacy and influence global affairs. Although, as he writes: "no one has yet ventured to estimate what the benefit to the world economy as a whole may have been; but that it was a benefit and not a cost seems beyond dispute" (Ferguson 367)

[7] Robinson and Gallagher argue that policy-makers in South Asia "were determined to see off any challenge to British 'paramountcy', party in order to protect trade links […] policy-makers in London generally fell in with their wishes" (Potter 28)

of the officials who manage the wealth that they set up systems of communication that both allow for a transformation in the scale of human organisation as well as reducing the potential conflict it may create. The economic aspect of colonialism brought with it a need for a bureaucratic system that could manage the population using values within an imposed system they understood. The census provided a method of discovering the scope of the type of system needed by the government to ensure that the economy was of performing well.

The early modern globalisation phenomena that was marked by transcontinental exchange was also characterised by the sharing of quantitative information for the express purpose of imposing a structure on society. Herein, the role of technology as a means of gathering and processing large amounts of information played a central role. Information that can be assessed by a multitude of institutions, irrespective of location or knowledge of the area wherefrom the information originates, has a value that can be reduced to correlations and patterns. Sharing information that has been reduced to numeral values that can be filtered and adjusted depending on the requirement is more easily interpreted. Cohn, Said and Foucault argue that the "official" construction of societal paradigms are the product of classifications rendering heterogeneous entities into homogenous ones for purposes of managing the society. Sharing information in the early modern period was done within the British Empire on the basis of constructing a bureaucratic system capable of managing the population present in India. The British thus systematically classified components of Indian society into areas that could be managed by the government to a certain end. The information and knowledge that was gathered and reduced to quantifiable data did not necessitate the need for knowledge on particular characteristics that made up different groups. This led to the continued practice of keeping the caste system in place. Once the system was set in place its functionality could be shared throughout the empire. It is one of the key points that N. Ferguson makes when he argues that the British Empire developed a "global network of modern communications. It spread and

enforced the rule of law over vast areas [and] maintained a global peace" (Ferguson 366). The progression of globalisation of expanding scale and magnitude brought with it the development of technology and thereby the quantity of information. The method of collecting and building knowledge on a society may have changed in scale and magnitude with the advent of the big data systems and in particular the Aadhaar card. However, the rationality that makes use of these methods and the implications on societal change are continuations of earlier methods used since the early modern period. Jacobson enforces this point with the development of the UIDAI and transnational sharing of information when arguing that "at the heart of this transnationalized practice is the role of technology as a means of increasing economic productivity and welfare" (Jacobsen 464).

Governmentality

The development of the caste system as an institution represents how the British imposed a structural classification on society. The rationality of the government determined the classifications of society based on the census. The census in this way represents a new technique of the government in order to gain insight into the economics of the population. In an attempt to improve the economy of the population the government, or official mind, systematically segregated groups of people and imposed a structure that would enable a more efficient management of the population. There was a superimposed structure determined by the British that assumed a central power position during colonialism. Upon the dissolution of the colonial structure the central power was handed to Indian congress. However, the continuation of the rationality that structurally segregated portions of the population remained. Chatterjee discusses how B. R. Ambedkar "had to contend with the fact of governmental classification" (Chatterjee 13). The institution put in place by the British continued. More importantly however, the rationality that maintained the need for a social structure continued. The techniques that the British used were imposed on Indian society and later adopted within the framework of the government. It is a sterling example of the rationality of government that Foucault developed. The development of technologies for gaining quantitative data on the population without requiring the active participation of that population, as Chatterjee suggests, are "'pastoral' functions of government […] using similar governmental technologies all over the world but largely independent of considerations of active participation by citizens in the sovereignty of the state" (Chatterjee 47). The Aadhaar card and the UIDAI collecting data on the population, with the goal of gaining complete registration by every Indian citizen, is a technique or technology developed by the government. The government may have moved beyond the structural classification of society by developing a system that no longer requires segmentation to create homogenous entities because it is now able to group the entire

population within indicators that are now proving important for the welfare of the population. It is here that Foucault's governmentality is important.

Foucault developed 'Governmentality' in a lecture in 1978 at the Collège de France. Governmentality is the rationality of a government or art of government which he says "is essentially concerned with answering the question of how to introduce economy – that is to say, the correct manner of managing individuals, goods and wealth within the family and of making the family fortunes prosper" (Foucault, Governmentality 92). Although Foucault begins with a description of sovereignty as a transcendent power he is primarily discussing the changes undergone as the population has increased and governmental structures have changed. What began as a transcendent sovereignty of the ruling power has become a more integrated means of managing the population by government. Foucault acknowledges that there have been changes since the initial 'transcendence' of state to manage the economy. This development is what Foucault discusses in relation to new technology. Here he describes that "through the development of the science of government[8] […] it became possible to identify problems specific to the population […] and that 'statistics' […] now becomes the major technical factor, or one of the major technical factors, of this new technology" (Foucault, Governmentality 99). Foucault argues that the adoption of new technology by the state with the goal of assessing the problems of the population has three consequences. First, the family becomes an instrument in relaying information regarding the population and its issues back to the state. Second, the welfare of the population becomes the goal of the government. Finally, the science of government becomes a series of techniques by which the population is managed and from this comes "the birth of political economy" (Foucault, Governmentality 101). As the welfare of the population becomes the goal of the government through managing the economy of the

[8] Foucault describes this as a result of the emerging problem with growing population and wealth which resulted in the government "recentring the theme of economy on a different plane from that of the family" through the "science of government" (Foucault, Governmentality 99) which deals with the economics of the population (rather than the family) through statistics. He calls statistics the 'science of the state' (*ibid.* 96).

population, the instruments used to gather information on the population take precedence rather than the notion of sovereignty. He himself argues that "this is the new target and the fundamental instrument of the government of population: the birth of a new art, or […] a range of absolutely new tactics and techniques" (Foucault, Governmentality 100). The negotiation between the state, civil society and the population under this 'new technology' paradigm becomes increasingly one that is dependent on the information provided by each in order for the rationality of the government to improve the welfare of the population. Contemporaneously, rather than a transcendent sovereignty there are a variety of institutions that relay information back to the government. The tools the government uses to assess the welfare of the population are now carried out through and in collaboration with the institutions that the population have access to.

Institutions act as government or civil society established structures that regulate the activity of various components of society. F. Fukuyama argues that "[Institutions] are, in essence, persistent rules that shape, limit and channel human behaviour" (Fukuyama 6). By comparison, D. North argues that "Institutions are the humanly devised constraints that structure political, economic and social interaction" (North 97). Despite there being an overlapping emphasis on the role of institutions between the two authors there is a disparity in institutions across different countries. L. Dumont argues that the caste system in India is a form of institution that is sustained by a social principle, namely, hierarchy. Using the definition for hierarchy proposed by T. Parsons he argues that "our modern denial of hierarchy is what chiefly hinders us in understanding the caste system" (Dumont 20). Essentially, Dumont is arguing that the allocation of value and evaluation were notions imposed on Indian society by the British but which no longer exist in similar forms in parts of Europe. This leads to a disjunction between the kind of institutions present in the West and the ability to understand how institutions function differently between India and Britain, despite being formed out of similar

processes in the West. He therefore suggests that new ways of interpretation and understanding are necessary. In order to better understand the functioning of institutions within a governmentality perspective, C. Tilly offers a means of assessing the viability of institutions by looking specifically at urban environments. Tilly argues that "cities offer privileged sites for study of the interaction between large social processes and routines of local life" (Tilly 704). As such, the operation of the UIDAI in a city offer a basis for ascertaining the interactive role between a big data institution and the population. Moreover, the kind of institutions present in a city represent how the government is able to regulate and maintain the population.

L. Lucassen outlines a schematic framework for assessing what kind of institutions are available to populations in cities. He describes five[9] models, these are the: full citizen mode; with high amount of institutional support for population, the ethno-national model; with institutions that segregate ethnicities and populations based on their religion, external differential citizenship model; with institutions that exclude foreign migrants, internal differential citizenship mode; with institutions that exclude internal migrants, 'empty citizenship' model; with almost no institutional support for a population (Lucassen 665). The type of interaction that exists between the state and its view on the people within its territory largely determines the type of institutional form that exists there. More importantly however, the existence of citizens and non-citizens – albeit, for example, because institutions do not recognise the people as such or because the people themselves do not make their presence known to the institutions – is an important factor in the actions made by the government in line with their rationality. As governmental institutions are components that monitor the welfare of the population they thus both provide and collect information on the population. Those people whom fall outside of the periphery of institutional support, because, for instance the state

[9] Although he tentatively acknowledges a sixth as a 'lighter' version of the 'empty citizenship model' (Lucassen 667).

employs an internal differential model or the institutions are stretched too thin to provide meaningful support to the population by virtue of its sheer size also fall outside of the relay of information between the institutions and the state. The existence of institutional support set up within the civil society offer a provisional support structure. In an attempt to improve health and literacy for children in urban slums the Indian government began the Integrated Child Development Scheme (ICDS). This represents how "the residents of [slums] could organize to get themselves identified as a distinct population group that could receive the benfits of a governmental program" (Chatterjee 56). This type of initiative also reflects how the government negotiates with the population. The means of negotiation becomes the 'civil society' which "mediates between the Many and the One" (Hardt and Negri 328). However, civil society is also an aspect of the state in that it "organized capitalist society under the order of the state and in turn spread state rule throughout society" (Hardt and Negri 328). The institutions set up to support urban areas that fall outside of the existing governmental scope serve to support the economy of the family. Secondly, when these methods have been adopted by the family, information is relayed back to the government. The type of institutions within the urban environment and those that the population makes use of provide the information that underline the rationality of the governments decisions.

The range of institutions that operate within a city, despite there being an overlapping emphasis on regulation, differ greatly but serve to mediate between the population and the government on different issues. Moreover, as Lucassen argues there are institutions that are specifically orientated toward exclusion of portions of the population. Thus, an institution regulates but does so differently depending on the government. Many governments have set up a number of institutions that operate to improve the government's performance, such as educational- or healthcare- facilities. The movement of people to these areas is often a result of opportunity being present in these cities and structures to support the population should ill

befall them. However, in cities that house large proportions of the country's population there may be a "systemic disjuncture between opportunity structures for livelihood" (Koonings and Kruijt 10). The result of a large population growth, according to Foucault, leads to the creation of more governmental techniques to monitor and affect the population's welfare. Yet, there exists an intermediary phase wherein the techniques in place do not benefit the welfare of the population. In contrast however, E. Glaeser argues that urban environments house institutions that are responsible for the success of a city, regardless of the current climate. He contends that the successful cities attract skilled workers by way of institutional support systems that implement "raw political power or sensible pro-business policies [or] have thrived by establishing themselves as bastions of economic freedom and the rule of law" (Glaeser 224). Institutions that are set up in order to regulate urban environments are often in close proximity to the population. This proximity is one of the core values that Tilly, Glaeser and other urban historian share when assessing urban environments. It is the contact to the population that enables institutions to react more quickly to the needs of population. In relation to governmentality the proximity of the institution allows it to gather information on the population quicker and thus enables it to adjust its techniques in order to improve the economy. The existence of a disjuncture between opportunity structure and livelihood in large cities is more quickly recognised by the institutions within that environment. How these institutions are ordered to deal with the problem is what marks the governmentality. The development of the ICDS is a form of inclusionary policy set up by the government. The development of NERPAP for voters is a form of exclusionary policy.

Institutions specifically orientated toward exclusion exist within the same rationality as that which seeks to benefit the welfare of the population. Foucault develops the idea of exclusion through a historical analysis of the creation of asylums. Here, he argues, there was initially a "religious segregation for purposes of moral purification […] to effect moral

syntheses […] by practicing a social segregation that would guarantee bourgeois morality and a universality of fact" (Foucault, Madness and Civilization 246). The exclusion of people from the population is a way to create a more homogenous group. The caste system set up large homogenous groups and extrapolated values based on the census. This represents a synthesis of exclusion and inclusion[10] in the governmentality. Institutions that make use of big data systems are able to include information that comes from a multitude of sources and therefore mark a new form of inclusionary methodology by the government. These big data institutions, like the UIDAI, set up under the governmentality paradigm can thus function as a means of regulating exclusionary policy in new ways when the welfare of the population is considered the end-goal. As the existence of exclusion falls under the rationality for improving the welfare of the population, albeit it in a moral sense or an economic one. The proximity of institutions in urban environments to the population allows them to quickly ascertain the problems of the population but from here the diversity of the population (albeit it in ethnicity, wealth, religion, etc.) could prove to be an issue in the face of the welfare of the population. The density of population was a reason for the formation of the science of government. In areas that are witnessing unprecedented concentration of population, and without new techniques in policy formulation for the political economy, the existing techniques of the government including exclusion are used to interpret issues and implement new policy. However, big data systems such as the Aadhaar card are providing the government with new and more inclusionary data sets that they can use to act on in new ways. Not merely by interpreting issues but also, with extensions made by the UIDAI to public health institutions, in predicting issues.

[10] As the system applied a value dependent on the census response which represent an inclusionary approach but at the same time made no distinctions within those groups thereby excluding those values of the community.

Information Governmentality

Identification Technology in the Modern Age

An important feature of the modern state is the ability for the government to identify the population within the territorial boundaries. Foucault argues that for the government to ensure the welfare of the population there are systems in place that relay the information from the population to the government. As the population increased the need and appropriation of technology and techniques that could properly approach the plethora of families within the territory become more intertwined. Implicit herein is the need for identifying who is within the boundaries of the state. Hildebrandt argues that "the construction of the territorial nation state required the identification of those aligned to the territory and the nation." (Hildebrandt 56). The technology available today enables the government to gather more information on the identities within the population. This is a direct result of a large portion of the global population becoming connected to digital platforms of exchange. As such, the amount of information that has become available on the population has increased. In combination with the increased processing and storage power made available by computer systems it has created Big Data which "refers to things one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value, in ways that change markets, organizations, the relationship between citizens and governments, and more" (Mayer-Schönberger and Cukier 6). Therefore, the new insights are no longer purely related to the identification. The census in India was used as a statistical tool to identify the population. Yet, aside from simply identifying the population, one of the effects was the production of knowledge about the population. Systems such as Aadhaar are tools to identify in order to assist governing, similar to the census, however a novel issue occurs with big data systems and identification. An "asymmetry of effective access to knowledge", as Hildebrandt aptly names it, which concerns the knowledge

being gained in big data systems, "this knowledge is protected as part of a trade secret or intellectual property the citizens to which this knowledge may be applied have no access whatsoever" (Hildebrandt 63). Chatterjee writes that "the contemporary regime of power […] secures legitimacy not by the participation of citizens in matters of state but by claiming to provide for the well-being of the population […] its apparatus is […] an elaborate network of surveillance" (Chatterjee 34). Through the network of surveillance that big data systems enable the government is able to deduce issues facing the population. However, identifying the population in the modern state has taken a new step that moves beyond identifying the population in order for the government to act. As the type and quantity of information the population provides increases the government is able to judge the effect of the environment on their behaviour.

One of the features that big data systems like the Aadhaar card enables the government to do is create biometric and demographic profiles which can aid the government when searching for issues. Profiling concerns the "interplay between monitoring and adaptation" (Hildebrandt 57) and is therefore an extension of identification. By including the relevancy of the environment the interplay between monitoring and adaptation becomes tied to monitoring and pattern recognition of actions taken to adapt. Profiling within a big data system has three features: "First, […] profiling by machines (Elmer 2004). […] these machines are software programs 'trained' to recover unexpected correlations in masses of data aggregated in large databases. Second, […] discovering knowledge we did not know to be 'hidden' in the data (Zarsky 2002–2003; Custers 2004). Thirdly, we cannot reflect upon the way that profiling impacts our actions because we have no access to the way they are produced and used" (Hildebrandt 58). These features are not exclusive to the government. Companies such as Google have made claims that one of the goals of gathering more information on the behaviour of individuals "is to enable Google users to be able to ask the question such as [...] 'What job

shall I take?' (Rubinstein, Lee and Schwartz 273). The government is able to use the big data systems in a similar fashion. Rubinstein *et al.* outline two key ways of searching through databases or data mining: Subject-based searches; these are means to enhance the search of a particular individual or group, and Pattern-based searches; these start with a theory or model of assumptions which then searches through a database finding patterns that match. The government uses the ability to construct profiles and search on the basis of established patterns or models of assumptions on the population to provide more efficient responses to the issues facing the population.

Incentives for using Big Data Systems

One of the important features of providing more efficient responses by the government is the reduction in cost. When the actions undertaken by the government become increasingly well-managed and structured in providing the necessary support, the cost decreases[11]. Initiatives described by Chatterjee such as ICDS show how the government is investing in initiatives that seek to improve the welfare of the population that currently fall outside of (full) institutional support in urban environments. The president stated in his speech in 2012 that the government is aiming to reduce corruption by "efficient and automated delivery of public services with minimum human intervention" (Press Information Bureau, Government of India). Essentially, this indicates how the government is beginning to perceive the benefit of using big data to change "the relationship between citizens and governments" (Mayer-Schönberger and Cukier 6) by reducing the cost of gaining information on the welfare of the population. Big data systems in this way offer the government an ability to develop policy in line with the needs of the population. Essentially, once Aadhaar cards have been integrated within prominent urban institutions such as educational facilities or medical facilities, the government receives a

[11] Cost here does not necessarily refer purely to the financial cost but also to any additional cost that the actions may incur such as corruption.

constant inflow of information on the standard of welfare of the population. Moreover, it is no longer only the registered citizens who are providing the information, the urban environment becomes able to relay information on the citizens to the government in real-time. The environment "becomes the interface and its capacity to perform real time monitoring and ubiquitous and proactive computing" (Hildebrandt 60). As the cost of maintaining the standard of welfare is reduced by the more efficient actions made available to the government, it can perform using the information from both the population and the environment to more efficiently target the population.

The benefits of having big data systems in place and a degree of automation in pattern recognition may allow a better response to the issues of the population. However, it is also important to recognise that the population becomes more obliged to provide all the data they can. In order for an automated system to produce the correct pattern of an issue for the government to target, as many variables as possible need to be recorded. Cukier and Schönberger argue that "when the quantity of data is vastly larger […] exactitude in some cases is no longer the goal so long as we can divine the general trend" (Mayer-Schönberger and Cukier 40-41). Under the governmentality paradigm the ability for the government to monitor and gather information on the population is a central feature in their ability to respond. That the UIDAI is attempting to link to other institutions is an attempt to further integrate big data systems in institutions. This is also a feature of institutions becoming reliant on these systems Moreover, it also marks a change in how reliant institutions are becoming on the data itself rather than the system's ability alone. Institutions become better able to document the issues facing the population with systems that register the behaviour of the population. The proximity to public services plays an important role here. As the quantity of data increases, more accurate patterns become apparent. Essentially, governments begin to map out the issues the population will come to face based on the experience of people who encountered issues before others.

Issues of privacy and autonomy become apparent here. When the institutions become increasingly dependent on big data systems, the way that citizens interact with the government begin to change. Where Tilly argued that "citizenship designates a set of mutually enforceable claims relating categories of persons to agents of governments" (Hanagan and Tilly 253) the reliance on data and the systems that search and compile patterns did not exist to the extent they are employed today. Chatterjee claims that there is a new distinction between citizens and populations being made by institutions where "citizens inhabit the domain of theory, populations the domain of policy" (Chatterjee 34). Essentially, citizenship is seeing a marginalisation of the mutually enforceable claims. The institutions capable of monitoring the activity of the population need less interaction from the citizens in order to view the issues facing the population. Rather, these institutions are moving beyond the capability of responding to the claims of the citizens and into the realm of policy capable of predictive exclusion.

Big Data and Exclusion

When IBM pitched an idea[12] about using smart environments and big data systems to improve the current education system it reflects strongly the issue that Balibar presented when talking about segregation in education in a techno-political society. Balibar argues that "technological transformations will assign educational inequalities and intellectual hierarchies an increasingly important role in the class struggle within the perspective of a generalized techno-political selection of individuals" (Balibar and Wallerstein 26). Big data systems and pattern-based results can potentially lead to increased segregation of the population. What IBM is essentially proposing is that the educational institutions adopt a pattern-recognition system

[12] "Within the education realm, predictive analytics can also help a school district's challenge with making static, point-in-time decisions about programming and failing to consider the dynamic needs of students, which results in wasted resources and sub-par student performance. IBM SPSS Student Performance solutions use advanced algorithms to adapt and match the right programs to each student throughout the year, improving standardized test scores, preempting dropouts, and creating a positive, customized experience for each student" (Morelli, Shearer and Buecker 19)

that takes into account the data provided by an individual. Autonomy and privacy, the barriers between the population and the government, fade slowly as institutions and companies gain insight into the issues facing the population from both the environment and the individual. As such, there is an influential risk when a disjunction between the citizens who select which information is given to the government and the information that the government requires occurs. "Precisely because a person is not aware of the profiles that are applied to her, she may be seduced to act in ways she would not have chosen otherwise" (Hildebrandt 63). The influence that the government or a company like Google can exert on the activities of the population can affect the autonomy of the individual. However, the individual may have submitted information that is categorised into on profile that is not applicable. The so-called "false positive occurs when a data relationship identifies an innocent individual" (Rubinstein, Lee and Schwartz 260). Alternatively, the opposite may be true too. An individual may have a positive relationship within the pattern but for lack of information is excluded. Segregation and exclusion in pattern based systems is the result of false, incomplete or consciously omitted data in the databases and profiles.

As more items become datafied, and as big data systems become more integrated in institutions, the policy of the government may come to reflect the informational patterns by big data systems. "Datafication" is how K. Cukier and V. Mayer Schönberger see an important step being made. They argue that "big data is also characterized by the ability to render into data many aspects of the world that have never been quantified before; call it 'datafication' (Cukier and Mayer-Schoenberger 29). However, datafication is a process that marks an important inefficiency in pattern-based decisions. It refers to items that have not been numerically recorded yet, thus items that cannot be correlated within a pattern and implicitly acknowledges the shortcomings of policy based on data correlations. M. DeRosa worries that "there will be great temptation for the government to … take action based on the results of

data-analysis queries alone" (Rubinstein, Lee and Schwartz 269). Although a big data system and the means with which knowledge is gained from it can provide efficient policy recommendations, it also contains acknowledged short-comings that may lead to exclusion of the population. The Aadhaar card provides a means of identification and is becoming an important tool for the government to identify the population and for the citizen to access particular public services. As technology improves and the information relayed between citizen and government increases, as a result of datafication, the ability for the government to improve the welfare of the population, that is registered, increases. However, it is also a reflection on a segregated institution or an institutions that has the potential to exclude. Essentially, big data systems in this way become a tool of the bureaucratic official mind in better sorting which groups of the population should be dealt with and in what way. An individual or group may begin to see different treatment or social classification based on the information provided. As such, exclusion is still a feature of institutions that adopt big data systems. These systems provide the government with an enhanced ability to monitor the welfare of the population. However, on the basis of this it creates a system that can potentially segregate and classify more people than the caste system.

Big data systems are, theoretically, providing a basis for the government to improve the welfare of the population. It is also shown that there is an aspect of segregation in the institutions that look at the population. Exclusion in a big data system, by not being registered in data points or profiles that institutions look at becomes very similar to being profiled in a category that you are not part of. E. Schmidt and J. Cohen write that one of the fears of having an e-government where the government monitors the population is that it is also able to monitor which people or who have programmes that redirect their identity and hide it (Schmidt and Cohen 94-95). People who do not register or give relevant data to the government, such as not providing demographic information when registering for the Aadhaar card, are not entitled to

the card. If the Aadhaar card becomes a feature of public service access then the not registering limits access to the institutions. Therefore, not only is there exclusion from access there is another consequence in that exclusion is discouraged by restricting access without participation. Moreover, current estimates claim that 67% of the population of India has registered to receive the Aadhaar card and almost 400 million people who aren't registered. When the caste system was first set up the census provided a tool for the government to categorise the population and impose a structure. Those people who did not fall into a predetermined category were still part of the population and the ensuing policy and structure imposed on society. Today the Aadhaar card allows the government to create more efficient plans by establishing profiles, performing pattern-based searches and identifying issues the population is facing. The 400 million people who fall outside of the big data system do not fall outside of the policy the government imposes for the whole population.

Conclusion

Summary

This thesis sought to investigate the role of big data in institutional exclusion by drawing on the Aadhaar card and the UIDAI as a case study. The Aadhaar card is a reflection on the new technology made available to organisations. Its role in politics in emphasized by move the government is making in automating services in order to reduce corruption. The role of identifying the population is framed within the governmentality paradigm by Foucault. Taking as a starting point the how the census functioned in identifying the population. The development of technology allowed more encompassing structures for identifying the population. Where the census provided the British government with a means of categorising society, the Aadhaar card provides the current government with a means of overcoming the earlier values attached to identity. The underlying themes in this thesis were concerned with tools for identifying the population and the degree to which the population can be excluded from the modern institutions. The progression of the government viewing the population as the area where policy is relevant by means of improved surveillance technology is resulting in institutions further integrating big data systems. However, as a result of this, the dependency of the information being provided becomes more important. Thus, as the UIDAI attempts to link to more institutions it makes the Aadhaar card a vital form of identification that may grant citizens access to public services. Furthermore, it is collecting more data in order to construct more accurate governmental indicators.

Final Argument

The extent to which big data systems in institutions allow the exclusion of the population can be answered in three ways. First, the individual today has access to technology that provides data for big data systems. The individual is able to enrol in a local UIDAI agency

and register for the Aadhaar card. In doing so the individual contributes to the database that the government can access and in turn is provided with a legitimate form of identification that is accepted widely in India. Upon gaining the card the individual is granted access to certain public services. The government in turn is able to identify the individuals within a population but is also able to assess the welfare of the population. It is in the interest of the government to gather as much information on the population as possible. This will allow it to produce more informed policy that targets the population better. On the other hand it is in the interest of the individual to provide the government with information for two reasons. First, as is the case with the Aadhaar card it provides access to public services. Second, it allows the government to better assess the needs of the population. The notion of incorporating exclusionary policy within institutions that contain big data systems is therefore restricted. This is because the government is able to more efficiently address the issues of the population and because the data is no longer only provided by statistics.

The second way the question can be answered is by the environmental factor. The urban environment is one of the areas that can be datafied. The interaction between environment and individual is recorded, albeit through recording financial transactions or satellite imagery. The ability for the government to continue with exclusionary policy in an environment that is potentially providing information on the population as much as any governmental institution is near impossible. However, the interaction between government and population changes because the issues facing the population are known to the government regardless of individual or group activity. The environment becoming smart enough to record and provide the government with data on the population no longer requires their active participation.

Third, an institution's ability to process the information and provide recommendations is increasingly automated. As the machine becomes a necessary tool for gathering and processing information there is no need for the government to impose cultural considerations

in its programming. A fully automated big data system that can improve the welfare of the population is linked to four points: (1) Requires registration within a big data system that the government uses or has access to. (2) Privacy regulations that the government imposes on the scope of information permitted to be assessed and put in a profile. (3) Autonomy of citizen within the big data environment means that not all information is made available by choice. (4) The tools by which the information is recorded in the environment and personal technology are already machines but the lack thereof influences the extent to which an automated database can provide patterns.

The extent to which governmental institutions allow exclusion is restricted by the degree of information provided by the citizen, the environment or the variables the system has access to. As such, rather than exclusion being a deliberate act by institutions or government, it is the extent to which the individual adopts the tools that relay information back to the government. This does not mean that exclusion does not continue. In fact, exclusion may take on a different form entirely. The ability for the government to gain information on the population and develop models of assumptions that search through the information may create a form of segregation. As exclusion under the Foucauldian paradigm is concerned with homogenising the population, regardless of the ability to collect data, the government may develop policy that seeks to create new homogenous categories. With a big data system in education such a homogenous category may be made up of students who performed in the top 10 percentile. Governments that integrate big data with institutions must remember that exclusion of portions of the population continues. However, exclusion is determined on the one hand by the information the government has collected and the assumptions it has made on the basis of that information. On the other hand it is determined by the information willingly provided by the individual, despite the ability for governments to influence their behaviour without them knowing.

Works Cited

Balibar, Etienne and Immanuel Wallerstein. *Race, Nation, Class; Ambiguous Identities*. New York: Verso, 2005. Print.

Ballantyne, Tony. "Colonial Knowledge." Stockwell, Sarah. *The British Empire*. Oxford: Blackwell Publishing Ltd., 2008. 177-197. Print.

Chatterjee, Partha. *The Politics of the Governed* . New York: Columbia University Press, 2004. Print.

Cohn, Bernard S. "Notes on the History of the Study of Indian Society and Culture." Singer, Milton and Bernard S. Cohn. *Structure and Change in Indian Society*. Chicago: Aldine Publishing Company, 1968. 3-28. Print.

Cukier, Kenneth and Viktor Mayer-Schoenberger. "The Rise of Big Data." *Foreign Affairs, 92:3* (2013): 28-40. Print.

Dirks, Nicholas B. "Castes of Mind." *Representations, 37* (1992): 56-78. Journal.

Dumont, Louis. *Homo Hierarchicus*. London: Cox & Wyman Ltd., 1970. Print.

Ferguson, Niall. *Empire; How Britain Made the Modern World*. London: Penguin Books, 2004. Print.

Foucault, M. "Governmentality." Burchell, Graham, Colin Gordon and Peter Miller. *The Foucault Effect; Studies in Governmentality*. Chicago: The University of Chicago Press, 1991. 87-104. Print.

—. *Madness and Civilization*. Abingdon: Routledge Classics, 2001. Print.

—. *The Order of Things*. New York: Routledge Classics, 2002. Print.

Fukuyama, Francis. *Political Order and Political Decay*. London: Profile Books Ltd., 2014. Print.

Glaeser, Edward. *Triumph of the City*. London: Pan Macmillan , 2011. Print.

Hanagan, Michael and Charles Tilly. *Extending Citizenship, Reconfiguring States*. Lanham: Rowman & Littlefield Publishers, Inc., 1999. Print.

Hardt, Michael and Antonio Negri. *Empire*. London: Harvard University Press, 2001. Print.

Hildebrandt, Mireille. "Profiling and the Rule of Law." *Identity Journal Limited* (2008): 55-70. Journal.

Jacobsen, Elida K. U. "Unique Identification: Inclusion and surveillance in the Indian biometric assemblage." *Security Dialogue, 43:5* (2012): 457-474. Journal.

Koonings, Kees and Dirk Kruijt. "The Rise of Megacities and the Urbanization of Informality, Exclusion and Violence." Koonings, Kees and Dirk Kruijt. *Megacities; The Politics of Urban Exclusion and Violence in the Global South*. London: Zed Books, 2009. 8-26. Print.

Lucassen, Leo. “Population and Migration.” Clark, Peter. *The Oxford Handbook of Cities in World History*. Oxford: Oxford University Press , 2013. 664-682. Print.

Mayer-Schönberger, Viktor and Kenneth Cukier. *Big Data*. London: John Murray , 2013. Print.

Morelli, Theresa, Colin Shearer and Axel Buecker. “IBM SPSS Predictive Analytics: Optimizing Decisions at the Point of Impact.” *IBM Redbooks* (2010): 1-58. Digital.

National Electrol Rolls Purification Authentication and Programme. “NERPAP; http://elections.tn.gov.in/NERPAP/.” n.d. *NERPAP - Information Pamphlet; http://elections.tn.gov.in/NERPAP/PDF/Pamlet.pdf.* 2 May 2015.

North, Douglass. “Institutions.” *Journal of Economic Perspectives, 5:1* (1991): 97-112. Print.

Potter, Simon. *British Imperial History*. London: Palgrave, 2015. Print.

Press Information Bureau, Government of India. *President’s Address to the Parliament 2012; http://www.pib.nic.in/newsite/erelease.aspx?relid=80784*. 12 March 2012. Website. 28 April 2015.

Rubinstein, Ira S., Ronald D. Lee and Paul M. Schwartz. “Data mining and internet profiling: Emerging regulatory and technological approaches.” *The University of Chicago Law Review, 75:1* (2008): 261-285. Journal.

Said, Edward W. *Orientalism*. London: Penguin Classics, 2003. Print.

Schmidt, Eric and Jared Cohen. *The New Digital Age*. London: John Murray, 2013. Print.

The Economic Times. *UIDAI covers 67% residents, generates 81.78 crore Aadhaar numbers*. 24 April 2015. Website. 2 May 2015.

Tilly, Charles. “What Good is Urban History?” *Journal of Urban History, 22:6* (1996): 702-719. Print.

Unique Identification Authority of India. *UID and Education; https://uidai.gov.in/images/resource/UIDandEducation.pdf*. n.d. Website. 19 April 2015.

—. *UID and Public Health; https://uidai.gov.in/images/resource/UIDandPublicHealth.pdf*. n.d. Website. 19 April 2015.

Wickramasinghe, Nira. “Colonial Governmentality: Critical Notes from a Perspective of South Asian Studies.” (n.d.): 32-40. Print.