Big Data and the Illusion of Choice: Comparing the Evolution of India's *Aadhaar* and China's *Social Credit System* as Technosocial Discourses

Saif Shahin and Pei Zheng

Saif Shahin
American University
School of Communication
4400 Massachusetts Avenue NW
Washington, DC 20016
(512)299-6145
saif.shahin@gmail.com

Big Data and the Illusion of Choice: Comparing the Evolution of India's *Aadhaar* and China's *Social Credit System* as Technosocial Discourses

ABSTRACT

India and China have launched enormous projects aimed at collecting vital personal information regarding their billion-plus populations — and building the world's biggest datasets in the process. However, both Aadhaar in India and the Social Credit System in China are controversial and raise a plethora of political and ethical concerns. The governments claim that participation in these projects is voluntary, even as they link vital services to citizens registering with these projects. In this study, we analyze how the news media in India and China — crucial data intermediaries that shape public perceptions on data and technological practices — framed these projects since their inception. LDA topic modeling suggests news coverage in both nations disregards the public interest and focuses largely on how businesses can benefit from them. The media, institutionally and ideologically linked with governments and corporations, show little concern with violations of privacy and mass surveillance that these projects could lead to. We argue that this renders individuals structurally incapable of making a meaningful "choice" about whether or not to participate in such projects. Implications for various stakeholders are discussed.

Nandan Nilekani, a technology entrepreneur who famously coined the phrase, "The world is flat" (Wired, 2005), used to head Aadhaar, the Indian government's project to build a database containing 1.3 billion citizens' demographic and biometric information — including fingerprints and iris scans. The project has been described as "the most sophisticated ID program" in the world by World Bank chief economist Paul Roemer (Bloomberg, 2017). Facing a challenge in court over concerns that Aadhaar would imperil Indian citizens' privacy, Nilekani penned a newspaper column in which he justified the collection of personal data on the grounds of "freedom of individual choice." He argued that "enrolment in Aadhaar is voluntary and individuals granting permission for [data collection] for their own convenience and benefit hardly qualifies as a violation of their right to privacy" (Nilekani, 2015, para. 4).

Nilekani's argument was later described by a petitioner as "disingenuous" because individuals really had little choice but to enroll if the government would oblige them to have Aadhaar IDs for availing even basic social and financial services (Chandrasekhar, 2015). And yet, there was more than barefaced disingenuity to Nilekani's invocation of the "freedom of individual choice" as the legal and moral basis for Aadhaar. It also reflected an ideological belief that individuals made informed cost-benefit choices *on their own* and were therefore solely responsible for them. This ideological belief underpins virtually all user agreements employed by technology companies that directly or indirectly collect users' personal data — such as Google and Facebook (Hoofnagle & Urban, 2014). As a result,

individuals — instead of governments or corporations — become what Baruh and Popescu called the "locus of privacy decision" in such matters (2017, p. 585). All forms of personal data collection, along with their (mis)uses by governments and corporations, come to be justified in the name of individual choice.

Surely, expecting citizens to make such choices and holding them responsible for their choices presumes that they fully understand the costs and benefits of those choices. It further assumes that the social and cultural institutions that shape public understandings of technology services and data projects — what Sawicki and Craig (1996) called "data intermediaries" — transparently present their costs and benefits to citizens. The news media, the most common sources for citizens to learn about government-run data projects, are perhaps the most important data intermediary shaping public perceptions of data projects and their potential impact on private lives (Schrock & Shaffer, 2017; Shahin, 2017). As Quail and Larabie observed, "media discourses reflect wider socio-political values regarding the public interest and public utilities, the relationship between the news and an informed society, and most specifically, the mobilization of ideology and power of naming and circulating narratives and truths" (2010, p. 38). Indeed, this was probably the reason why Nilekani himself chose to write a newspaper column to contest the petition against Aadhaar.

Proceeding from this premise, our study focuses on understanding how the news media covered two of the world's biggest government-run data projects: Aadhaar in India and the Social Credit System in China. The latter collects information on citizens' consumer activities and social behavior to predict their chances of paying credit card bills and mortgages. Both these projects exemplify the widespread belief that Big Data — referring to our growing capacity to gather, store, analyze, and cross-reference enormous quantities of digital information — can bring an end to "market crashes, ethnic and religious violence, political stalemates, widespread corruption, and dangerous concentrations of power" (Pentland, 2014, p. 16). Others have, however, pointed out a slew of ethical and political challenges that Big Data's emergence has created, including the "dangerous concentration of power" in the hands in governments and corporations that compile and control Big Data — especially in the aftermath of Edward Snowden's revelations that the U.S. government was colluding with technology companies to run surveillance programs on a global scale (boyd & Crawford, 2012; Shahin, 2016a).

Our study aims to unravel how Aadhaar and the Social Credit System have been framed since the Indian and Chinese media started covering them — in 2009 and 2007, respectively. In particular, we are interested in exploring the extent to which the news media made citizens aware of how these projects could compromise their privacy, leave them susceptible to surveillance, and further entrench the power of governments and corporations to determine how people lead their lives. We take a comparative and longitudinal approach that traces how media coverage differed between the two nations and evolved over the years within each nation.

Our analysis has both ethical and political dimensions. The legal supposition that consumers make individual choices about using Big Data-enabled technology services — and are therefore responsible for them — is ethically built on the conceptualization of citizens as "privacy pragmatists" who willingly give up their privacy for enjoying the

benefits and conveniences that technologies offer (Kumaraguru & Cranor, 2005). This belief persists even though empirical studies have showed it to be unfounded (Elueze & Quan-Haase, forthcoming; Hoofnagle & Urban, 2014). We argue that the broader socio-cultural milieu within which individuals make their "choice" — specifically the role of data intermediaries such as the news media — must also be a part of these conversations. In doing so, we draw attention to the structural relations between governments, corporations, and the news media and how they shape public understandings of Big Data and its impact and value for individuals and society.

SOCIOLOGY OF BIG DATA

A Technosocial Phenomenon

Big Data is typically viewed as a technological phenomenon of the information age. Three Vs are often invoked to describe Big Data — Volume, Velocity, and Variety, referring to the unprecedented scale of datasets, the unprecedented speeds at which data is being produced, and the unprecedented range of data types and forms (Ward & Barker, 2013). Such data is collected from a variety of sources, including sensor networks, government data holdings, credit cards, and public profiles on social networking sites. The three Vs focus on data magnitude, but equally important are the emerging techniques being employed to analyze and make sense of data (Andrejevic & Gates, 2014). Big Data analytics has found an enormous range of applications, from enabling corporations to predict customer preferences (McIntyre et al., 2015) to allowing scientists to tackle diseases (Mneney & Van Belle, 2016).

Beyond technology and analysis, boyd and Crawford identified a third crucial component of Big Data — mythology, or "the widespread belief that large data sets offer a higher form of intelligence" (2012, p. 663). This component draws attention to the techno-utopian view that Big Data are "objective" and untainted by human foibles (Pentland, 2014). The mythology of Big Data blinkers us to the fact that the algorithms that enable Big Data mining, storage, and analysis are ultimately produced by human beings and are subject to the same social forces that govern all aspects of human behavior. For instance, a Big Data tool for criminal risk assessment used in the United States was found to disproportionately predict Black defendants as criminals "at almost twice the rate as White defendants" (Angwin et al., 2016, para. 16). Racial bias against Blacks and Asians was also coded into facial recognition algorithms used by Google Photos and Nikon cameras (Crawford, 2016).

Privacy and Mass Surveillance

The collection and use — or misuse — of personal data by governments and commercial enterprises has long been a privacy concern (Garson, 1988). But the emergence of Big Data has amplified it manifold. Websites, social media technologies, telecom operators, credit card companies, and other data-based services collect users' personal data from a variety of data points — with or without users' knowledge. Even more insidiously, Big Data allows the cross-referencing of multiple datasets: this means even if one dataset does not have personal indicators, combining it with others may create them. For instance,

Acquisti and colleagues (2011) found that facial recognition applications can match anonymous profiles on dating sites with their public Facebook photographs. There are also examples of "de-identified" private information analyzed for commercial benefits (Ohm, 2009).

The legal framework within which governments and corporations operate often holds users themselves to be responsible for their privacy. Users are defined as "privacy pragmatists" who, while concerned with their privacy, are often willing to sacrifice it for enjoying the benefits offered by Big Data-oriented technology services (Kumaraguru & Cranor, 2005). Even authorities in the United States make it incumbent upon individuals to "evaluate their choices and take responsibility for the ones they make" (White House, 2012). Consequently it is individuals, instead of governments or corporations, who become the "locus of privacy decision" — a framework which "may not only fail to protect individual privacy, but also bias the privacy calculus of the larger society by reducing the level of privacy available to all" (Baruh & Popescu, 2017, p. 585). Belief in "privacy pragmatists" persists even though empirical studies have repeatedly showed that many people are often unaware of how technology services compromise their privacy and many citizens want more regulations on the activities of governments and corporations providing such services (Elueze & Quan-Haase, forthcoming; Hoofnagle & Urban, 2014).

India's privacy laws follow the same logic. For example, the Information Technology Act of 2000, a 2008 amendment to the act and the 2011 Rules define biometric data as "sensitive data that are only to be shared with consent" (Roy & Kalra, 2011). In other words, individuals are responsible for the security of their biometric data. But "consent" itself is so vaguely defined that it has little practical value (Dixon, 2017). In China, privacy has been defined as a fundamental right in the constitution (Stanaland & Lwin, 2013). In 2000, a few regulations were passed to protect internet information security and privacy. But the government has tightened its control over internet service providers in recent years, requiring them to record subscribers' telephone numbers and maintain the records for two months (State Council of China, 2000). In 2012, some social media platforms required users to provide real names to register. In 2016, the Internet Security Law of People's Republic of China was implemented, which asked all internet service providers solicit personal information when signing contracts with their users.

Closely related to the issue of digital privacy is the concern over mass surveillance by governments in collusion with technology companies. This practice became painfully evident when Edward Snowden, a U.S. intelligence employee, revealed in 2013 that the U.S. government was running global surveillance programs with the help of other governments and technology companies including Apple, Facebook, Google, Microsoft, Skype, Yahoo, and YouTube (Gellman & Poitras, 2013).

Safire (2002) argued that people curtail the extent of information they exchange over the internet, fearing that their online activities may be recorded and stored and possibly become accessible to government agencies for subsequent scrutiny. More recently, Stoycheff (2016) showed that Snowden's revelations had filled citizens with deep apprehensions about being constantly under surveillance and caused a "chilling effect" on the quality of online public discourse. In China and India, studies on surveillance have

focused on public health issues (Chadha et al., 2012; Shah & Mathur, 2010) — such as HIV (Sun et al., 2007), birth defects (Dai et al., 2011) and H1N1 (Liang et al., 2011).

BIG DATA PROJECTS IN INDIA AND CHINA

AADHAAR IN INDIA

The Indian government's Unique Identification (UID) project, known as Aadhaar (or "foundation" in Hindi), issues a 12-digit number to every resident, generated after recording their demographic as well as biometric data such as fingerprints and iris scans. The project got under way in 2009 and the Unique Identification Authority of India (UIDAI) started registering residents for Aadhaar in 2010. Media reports said nearly 99 percent of residents had been registered by the end of 2017. The government calls it "a strategic policy tool for social and financial inclusion, public sector delivery reforms, managing fiscal budgets, increase convenience and promote hassle-free people-centric governance" (UIDAI.gov, u.d. para. 4).

The project officially aims to curb corruption in the transfer of benefits to the public by taking out middlemen and allowing people, especially the poor, direct access to a range of welfare schemes. But despite claiming that participation was voluntary, the Indian government has made Aadhaar mandatory for citizens to access a variety of social and financial services — such as maintaining bank accounts and insurance policies, even for paying taxes. Still, even by 2017, there were hardly any legislative measures in place to protect the data being collected by the government and the technology companies that worked on its behalf (Dixon, 2017).

The project has raised fears of privacy violations and mass surveillance. Soon after it was launched, Indian intelligence officials said its purpose was to flush out illegal immigrants in the country — the reason why residents and not just citizens were asked to register for Aadhaar — and the claim that Aadhaar would help deliver welfare schemes was just a ruse (Tehelka, 2009). Security officials made attempts to use the Aadhaar database to find "criminals" — although they were stopped by the courts from doing so (Indian Express, 2014). There have also been instances of data collected under Aadhaar leaking out and becoming publicly available (Hindustan Times, 2017). Common citizens, however, mostly remain unaware of the dangers posed by Aadhaar to their privacy and its potential as a surveillance tool (Srinivasan et al., 2018).

SOCIAL CREDIT SYSTEM IN CHINA

Under the Social Credit System (SCS), the Chinese government proposes to assign a social credit rating to every citizen representing their "trustworthiness" — based on their everyday social and economic activities. The idea started being discussed in official and non-official circles in the early 2000s and draft planning outlines were proposed. The State Council of China's Planning Outline for the Construction of a Social Credit System (2014–2020), issued in 2014, states that the SCS will focus on four major areas: honesty in government affairs, commercial integrity, societal integrity, and judicial credibility. The goal is to "raise the awareness for integrity and the level of credibility within society" (China

Copyright and Media, 2014). To implement it, the government will track and evaluate "what you buy at the shops and online; where you are at any given time; who your friends are and how you interact with them; how many hours you spend watching content or playing video games; and what bills and taxes you pay (or not)" (Wired, 2017, para. 2).

In essence, SCS will allow the Chinese government, and the technology companies that work with it, to monitor every social, economic, and political action taken by every citizen — as well as most of their private activities. The proposal stems from a lack of social and economic trust that has engulfed Chinese society since its turn to market economics in the 1980s. Among other things, it prevents individuals and small businesses from getting loans from banks. But "rather than promoting the organic return of traditional morality to reduce the gulf of distrust, the Chinese government has preferred to invest its energy in technological fixes" (Hawkins, 2017, para. 5). The idea has also raised the hackles of data protection and privacy advocates, who have called it "an excuse to implement surveillance and control" (para. 17).

SOCIOLOGY OF NEWS

NEWS FRAMING

Media scholars have long argued that news is never an accurate reflection of an "out-there" reality. Instead, the news media "construct" reality by *framing* the coverage of issues and events in particular ways (Gamson & Modigliani, 1989). As Entman observed, journalists "select some aspects of a perceived reality and make them more salient in a communicating text, in such a way as to promote a particular problem definition, causal interpretation, moral evaluation, and/or treatment recommendation" (1993, p. 52). In other words, news coverage amplifies certain features of an issue or event, so as to create a certain perception of it, while marginalizing or disregarding other aspects. Framing coheres with the phenomenological perspective that "facts" don't have immanent meanings — instead, meanings are *made*. Those who control the meaning-making process are able to exert power by affecting others' perceptions, attitudes, and actions.

Several studies have showed a significant link between news frames and public perceptions of various political (Brewer & Gross, 2010) and economic (De Vreese, 2010) issues. But why do the news media frame issues in certain ways? Reese (2001) has argued that the framing of news is not arbitrary but follows a cultural logic. The way journalists cover an issue coheres with their ideological beliefs about the way the world is — or should be. These beliefs, in turn, are socially reproduced in and through news frames. Reese therefore defines frames as "organizing principles that are socially shared and persistent over time, that work symbolically to meaningfully structure the social world" (p. 11).

NEWS IDEOLOGIES

Culturally and critically oriented media scholars have identified various kinds of ideologies that produce news frames. Durham (1998), for instance, suggested that *positivism* was an ideology of journalism and deeply imbricated in the process of news framing. As an epistemology, positivism refers to knowledge based on empirical evidence.

As an ideology, positivism implies the deployment of "empirical evidence" to enforce particular interpretations upon facts and control the meaning-making process. Durham argued that journalists sought to control meaning-making by covering "atomistic events" in relation to other events across space or time — thus weaving them into specific social narratives and imposing particular meanings upon them.

A number of scholars have also emphasized *neoliberalism* as an ideological influence on news (Philo, 2014). While journalism is normatively expected to serve the public interest, it often ends up representing the interests of corporations and big businesses. In Western democracies such as the United States and Britain, this happens at least partly because news organizations are embedded within capitalist economic structures and rely on other corporations for advertising revenue, among other things (Herman & Chomsky, 1988). But the impact of neoliberalism on the news media has also been evident in nations such as India and China. Thussu (2007) rued the "Murdochization" of Indian television news while Shahin (2017) observed a general tendency in the Indian media to repose their faith in "benign capitalism." In a similar vein, Lee (2003, p. 18) accused the Chinese media of "wearing a socialist face with a capitalist body."

METHOD

The general objective of our study is to understand how the Indian and Chinese news media, as data intermediaries, framed Aadhaar and SCS respectively. We then use our empirical analysis to infer what kinds of ideologies shaped the news coverage and the implications of this social process for public understandings of these Big Data projects. To conduct our empirical analysis, we mined Indian news coverage of Aadhaar from 2009 — the year the project was launched — until 2016. A search in the Nexis Uni database for the keywords "Aadhaar" or "UIDAI" in the headlines or lead paragraphs of articles in Indian news publications in English provided a corpus of 5,090 articles (while India has a number of regional languages, English is the only language spoken nationwide; India also has an extensive network of English news publications; see Shahin, 2017). We also searched the Huike database of Chinese news articles for "social credit system." Coverage picked up in the year 2007 and yielded a corpus of 5,608 articles (in Mandarin) from that year until 2016.

Both these databases include a wide range of major and minor publications from the two nations. Our two corpora, therefore, represented a close approximation of the news coverage of the Aadhaar and SCS in India and China, respectively. But traditional qualitative or quantitative content analysis would have required us to draw smaller samples of a few hundred articles from these corpora — creating the possibility that we would leave out important elements of the coverage. We therefore conducted our analysis using *topic modeling*, a machine learning technique that enabled us to study both corpora in their entirety. Specifically, we adopted a form of topic modeling called Latent Dirichlet Allocation (LDA), which parses textual data to reveal the set of "topics" or themes that the articles focus on. Each topic is a set of "keywords" with a high probability of occurring proximate to each other across the corpus. Understanding the semantic links among the keywords reveals the meaning of the topic to the researcher. In addition, LDA also provides the proportion of use of each topic within a document. This allows the researcher to

understand which topics are more prominent in the text and which topics are marginalized. The technique is therefore useful for discovering frames in large volumes of news coverage. Because of its strengths, LDA topic modeling is being increasingly used for the analysis of news articles (DiMaggio et al., 2013; Jacobi et al., 2016). Comparative methodological studies have found it to be compatible with qualitative textual analysis (Shahin, 2016b) and more reliable than some other algorithmic research techniques (Guo et al., 2016).

For our study, we split the Indian and the Chinese corpora into yearly blocks to be able to track the year-by-year evolution of different topics in each corpus. Topic modeling was conducted using Python software. The package "genism" was used for analyzing the Indian corpus, while the Chinese corpus was analyzed using "jieba," a package that solves the issue of word split and text segmentation in Mandarin. The Python module for topic modeling includes stopwords (such as "is/are" "the/a") for the program to ignore. We also excluded words comprising more than 10 percent of an entire corpus to make sure that the keywords in our topics were not too generic.

RESULTS

NEWS COVERAGE OF AADHAAR IN INDIA

A five-topic model emerged from the Indian news corpus (see Table 1). Two topics dominated in the first couple of years, which we have labeled Biometric Data and Commercial Benefits respectively (see Figure 1). *Biometric Data* includes keywords such as "iris" and "fingerprints" along with "collected" — referring to the kind of personal data the government would collect under Aadhaar. This topic also includes "npr," "registrar," "census," and "enrollment" — keywords indicating where and how such data would be collected and stored (NPR is the abbreviation for National Population Registry). For instance, one article headlined, "Fingerprinting or iris scan during 2011 census for UID scheme" (Times of India, 2009), reported that the government would collect citizens' fingerprints and iris scans while conducting the regular census in 2011. Another keyword in this topic, "cabinet," refers to the cabinet of federal ministers who were creating the legislative framework for collecting biometric data. Overall, this topic signifies a focus on what kind of biometric data would be collected, how it would be collected and stored, and the legal/legislative framework that would enable this process.

The topic *Commercial Benefits* includes the keywords "business," "customer," "payment," "payments," "transactions," and "tax." In addition, the keyword "rbi" refers to the Reserve Bank of India, the country's federal bank, and "kyc" refers to Know Your Customer — a regulation that enjoins banks and other financial institutions to possess personal details of their clients. This topic relates to how commercial enterprises would benefit from Aadhaar. Several articles reported that Aadhaar would make it very easy for businesses to meet the KYC regulation. Others suggested it would simplify payments and transactions, as well as tax collection. A number of articles also talked about the Unique Identification Authority of India (UIDAI), the government body overseeing the project, working closely with business enterprises to facilitate data collection.

Table 1. Topic model of the coverage of Aadhaar in Indian news media.

Topic	Keywords
Biometric Data	npr, cabinet, iris, duplication, registrar, fingerprints, collected, census, enrollment, concerns
Commercial Benefits	digital, business, kyc, payments, transactions, rbi, customer, tax, payment, infosys
Administrative Process	post, enrollment, offices, ration, hyderabad, police, employees, administration, form, customers
Public Welfare	food, transfers, dbt, jharkhand, cent, pds, payments, enabled, payment, cylinders
Privacy Politics	privacy, bench, congress, voters, election, party, students, justice, electoral, bjp

Figure 1. Proportion of use of different topics in the coverage of Aadhaar (figures in %).
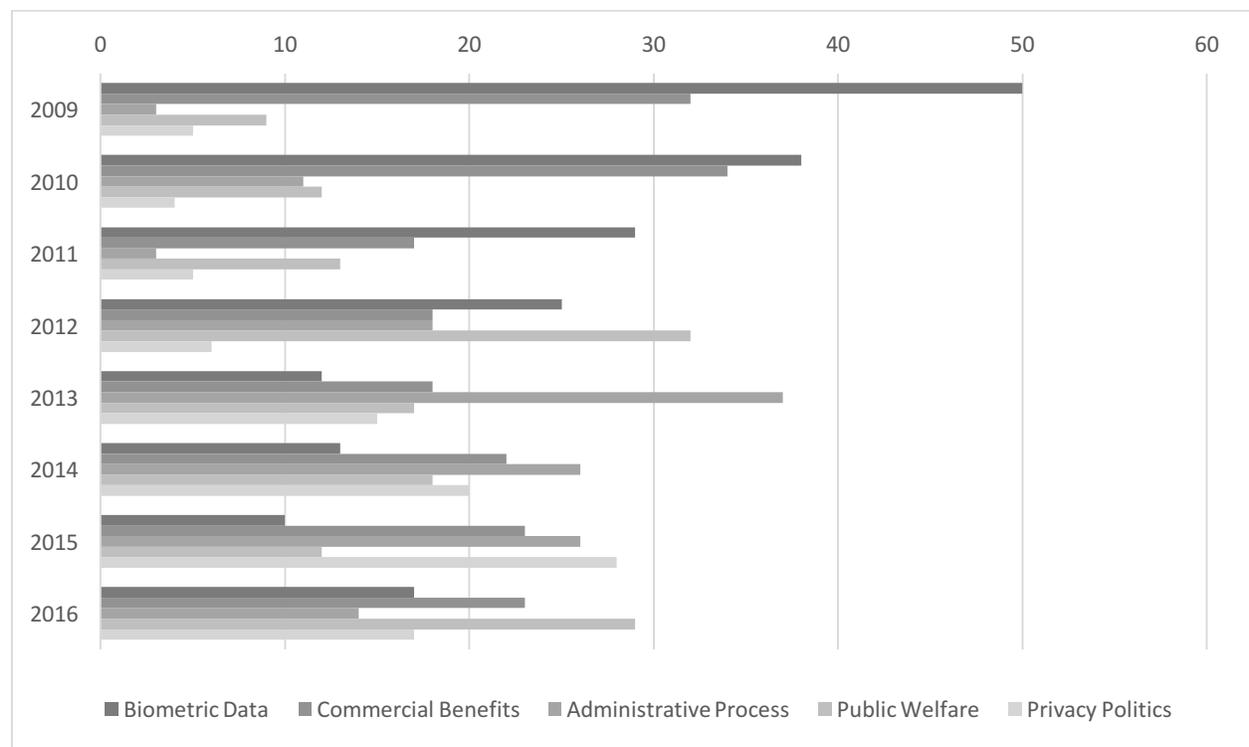
In 2011, a third topic— *Administrative Process* — became dominant the keywords "administration," "police," "post," "offices," "enrollment," "form," "customers," and "employees." This reflects a shift in focus of news coverage from the nature of data to be collected and the legal infrastructure that would enable it to the nitty-gritty of Aadhaar's implementation. Articles in this period often reported on how citizens were being enrolled for data collection and how administrative offices and services — such as police and post

offices — were being mobilized for it. As one article reported, "Currently, 14 post offices in the national capital are involved in enrolling citizens under Aadhaar which will be increased to 70 by the month end, the chief postmaster general (Delhi circle), Rameshwari Handa, said on Thursday" (Mail Today, 2011).

Yet another topic, *Public Welfare*, became dominant in 2012. With keywords such as "pds," "dbt," "food," "cylinders," "transfers," "payment," and "payments," this topic represents a concern with how common people, especially the poor, would benefit from Aadhaar. Several articles talked about the government being able to make payments and transfer various utilities to the public directly, without involving middlemen and thus reducing corruption. As one report said, "Aadhaar-enabled applications will be used for making pension payments, MNREGA payments, PDS distribution, scholarship payments, etc" (India Today, 2012). PDS refers to the Public Distribution System, using which the Indian government provides food and a few other essential items to the public at subsidized rates through fair price shops. The system has been beset for decades with problems such as middlemen siphoning off foodgrains and contractors at fair price shops charging customers unfair prices. Aadhaar, the media claimed, would circumvent these problems by allowing the government to transfer benefits directly to end-beneficiaries. The keyword "dbt" refers to Direct Benefit Transfer, another scheme launched in 2013 that formalizes the idea of direct transfer of government benefits to the public.

The topic Administrative Process once again dominated in 2013 and 2014, followed by Commercial Benefits. By this time, all the topics were being covered in reasonable proportions — including our fifth and last topic, *Privacy Politics*. This topic includes the keyword "privacy," "bench," and "justice." They signify the media taking note of litigations challenging the Aadhaar project over privacy concerns and courts getting involved in the matter. For instance, one article noted, "A Bench of Justices B S Chauhan and S A Bobde sought the states' views to ascertain their understanding of the nature of the UIDAI scheme and whether they had also linked Aadhaar cards to such services and if so the manner of doing it" (Indian Express, 2013). But while focusing on court proceedings, the articles rarely discussed the substantive basis of these litigations. The news coverage made little effort to inform the public about the specific ways in which these litigations claimed Aadhaar could imperil their privacy and leave them susceptible to surveillance.

Privacy Politics also included keywords like "congress," "bjp," "party," "election," "electoral," and "voters." The Congress is the oldest political party of India and was heading the federal government until the national election of 2014, after which its main rival, the BJP, came to power. The make-up of the topic suggests that privacy concerns were closely tied with this shift in Indian politics. Several articles reported that the new BJP-led government might not continue with the project, especially as a number of BJP members had criticized Aadhaar in previous years — including Narendra Modi, who became the prime minister in 2014. By 2015, though, it became clear that the BJP was going to push the project forward with a vengeance and make Aadhaar cards mandatory for citizens to avail of a range of public services. As one newspaper noted, "The Aadhaar project was started by the Congress-led UPA II government and is being defended by the Narendra Modi government. But the ruling BJP had assailed the scheme while in the Opposition"

(Telegraph, 2015). Privacy Politics was the biggest topic of news coverage in 2015, although only marginally larger than Administrative Process and Commercial Benefits.

NEWS COVERAGE OF SOCIAL CREDIT SYSTEM IN CHINA

The LDA analysis of the Chinese corpus reveals a model with five topics (see Table 2). A single topic, *Commercial Benefits & Rural Welfare,* dominates the coverage all through, with the year 2011 as the only exception (see Figure 2). Keywords in this topic include "small business," "rural residents," "rural area, "register," "bank," and "taxpayer." This topic relates to how businesses and citizens would benefit from SCS — especially small businesses and rural residents. These are two segments of the Chinese economy that have long struggled to borrow money from banks because they do not have as much capital as big companies and urban residents. The news media suggested that a more reliable credit evaluation system, such as the one offered by SCS, would allow them easier access to finance. Also, the SCS database would help the government identify impoverished rural areas more accurately and help them with welfare schemes more effectively. In addition, by rating tax payers and building a better tax management system, the SCS would enable conscientious tax payers to benefit from lower rates when borrowing from banks.
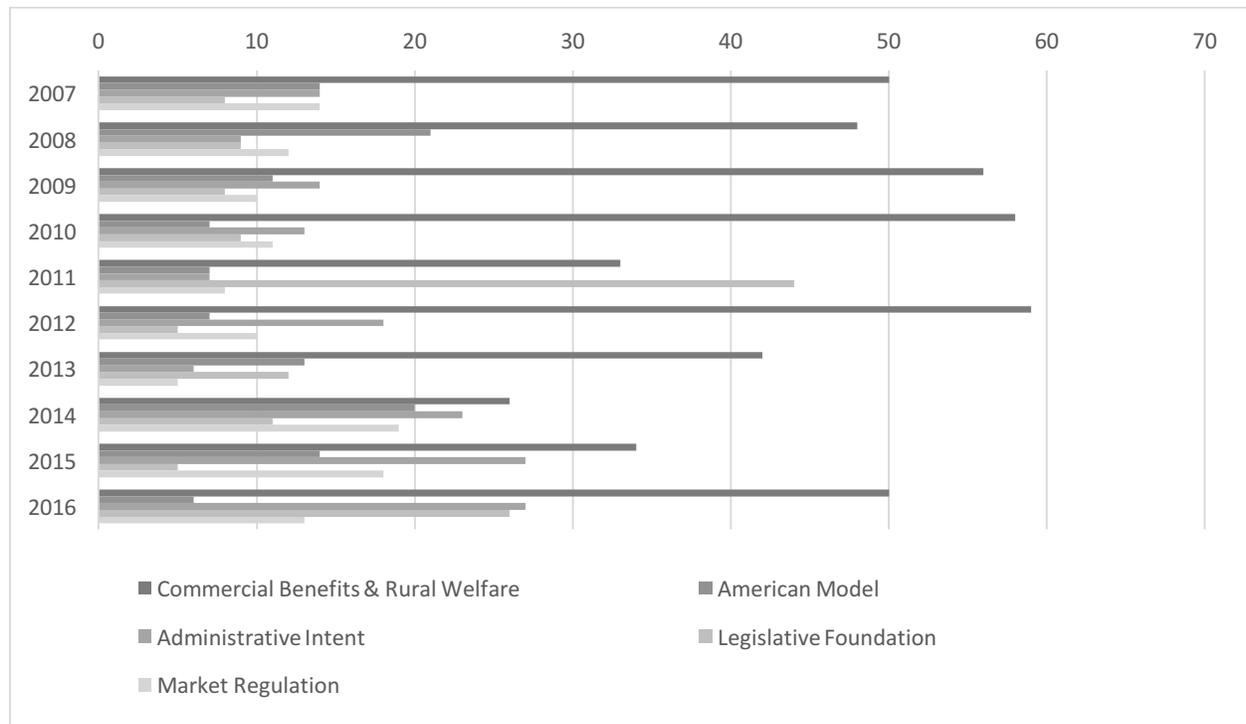
While Commercial Benefits & Rural Welfare dominated all through, other topics also gained secondary prominence at different stages. For instance, the topic *American Model,* with keywords such as "United States," "capitalism," "marketization," "should," "rules," and "trustworthy," was the second largest topic in the early years of 2007 and 2008, when SCS was being envisaged. It implies that from the outset, the Chinese media framed the idea of a credit system as not exclusive to China but one already in place in capitalist economies such as the United States. The media suggested that China "should" follow the American model to improve marketization and build a credit society where the rules of evaluating individuals were based on credit scores. This, however, was also the time when the subprime crisis hit the United States and exposed the limitations of a capitalist economic structure. But instead of changing their tune, the Chinese media framed the subprime crisis as "a lesson to learn in the importance of building up a credit system nationwide" (Zhou, 2009).

In the next two years, media coverage focused on how various levels of China's political and fiscal administration were determined to build the SCS. This is evident from the prominence of *Administrative Intent,* a topic with keywords such as "province," "province-level," "department-level," "each province," "People's Bank of China," "Pan Gongsheng (vice president of People's Bank of China)," "Lian Weiliang (deputy director of National Development and Reform Commission)," and so on. In 2011, *Legislative Foundation* suddenly became biggest topic, upstaging Commercial Benefits & Rural Welfare. Its keywords, such as "policy," "draft," "law," "legal," "illegal," and "hold (conference)," relate to the necessity and urgency of strengthening the judicial structure to prepare for the SCS, punish wrongdoers, and reward reliable individuals and companies. This was also year the when China's so-called "socialist judicial system" was completed, marking "a new stage in which there shall be laws to abide by, everyone should abide by the laws, the laws must be enforced strictly, and those who violate the law must be dealt with" (People's Daily, 2011). The media saw this system as laying the foundation for SCS.

Table 2. Topic model of the coverage of Social Credit System in Chinese news media.

Topic	Keywords
Commercial Benefits & Rural Welfare	small business, register, city government, bank branch, regulation, policy, shanghai, rural resident, pay tax, rural area, bank of china
American Model	united states, rules, expression, capital, trust, should, trustworthy, marketization
Administrative Intent	province-level, gongsheng pan, province, bank of china, each province, weiliang lian, department-level
Legislative Foundation	policy draft, law, phenomena, legal regulation, sexual, li keqiang, people's republic of china, hold
Market Regulation	professionals, price, internet, uncover, statistics, information safety, e-commerce, intellectual copyright, trust education

Figure 2. Proportion of use of different topics in the coverage of Social Credit System (figures in %).

But 2011 was an exception. From 2012 to 2016, Commercial Benefits & Rural Welfare once again became the biggest topic, followed by Administrative Process. Since 2014, however, the fifth and final topic — *Market Regulation* — also became quite sizeable. This topic includes keywords such as "internet," "e-commerce," "information safety," "intellectual copyright," "integrity education," "uncover," and "price." It reflects media

coverage of issues such as how to identify fake online transactions, strengthen online product quality inspection, inspect fake advertising and inaccurate online product descriptions, and bolster integrity education for citizens and business owners to build commercial integrity. The topic gained prominence when a Planning Outline for the Construction of a Social Credit System (2014–2020) was proposed in June 2014. This was also the time when Alibaba Group, China's biggest e-commerce and technology company with online sales and profits surpassing several major U.S. retailers (including Walmart, Amazon and eBay), made significant strides with moves such as offering its shares on the New York Stock Exchange. All these events led to the media coverage of SCS focusing on the need to build integrity and trust for encouraging e-business and e-finance in China. For example, an article quoted Jack Ma, the CEO of Alibaba, as saying, "I always believe that integrity can turn to money and it is the biggest wealth one can have…. Ten years ago, nobody dared to buy goods online, but now it is different because we trust each other. Trust, therefore, is the productive force" (Xinhua.net, 2014).

DISCUSSION

Juxtaposing the news coverage of Aadhaar in India and the SCS in China reveals a number of similarities and differences between them. The biggest difference is that no single topic dominates the coverage in India the way Commercial Benefits & Rural Welfare does in China. Also, the Indian media pay attention to the details of the biometric data to be collected under Aadhaar, especially in the early stages, while the Chinese media focus on China's structural readiness for a credit-oriented economic system. But such differences are to be expected. After all, not only are these two Big Data projects different in character and scope but the media and political systems in India and China are also vastly different (Shahin et al., 2016). India is a parliamentary democracy with a diverse range of mostly independent news organizations and publications (Chakravartty & Roy, 2013). China, in contrast, is a one-party autocracy where the news media are mostly state-controlled or self-censored (Fung, 2007; Pan, 2000). It is, therefore, the similarities in coverage that become more surprising and theoretically and normatively significant.

One common feature in both nations is that the media coverage is event-driven. This is problematic not because of what gets covered — but because of what doesn't. In India, where coverage began with the government's decision to launch Aadhaar, the media over the years took little note of the *absence* of a robust legislative framework that would clearly lay out Aadhaar's purpose and also define its limits (Dixon, 2017). The government kept expanding the project on a piecemeal basis by issuing decisions related to Aadhaar's administration and scope. These "events" — or what journalists call "news pegs" — were indeed covered. But as lawmakers were *not doing* anything to build a legislative foundation for Aadhaar, there were no news pegs for journalists to base their coverage on. The issue, significant as it was, slipped through the cracks of journalistic practice. In China, where coverage began when the idea of the SCS started being floated in official and non-official circles, the media's focus was on reporting how administrative and financial mandarins were positively campaigning for the project. Legislative Foundation did become a significant topic briefly when it was driven by an event: the 2011 launch of the so-called

"socialist judicial system" that was expected to form the basis for SCS. After that year, though, the issue was once again sidelined.

Second, even as Indian and Chinese media neglected to cover significant "non-events," major domestic and international events not related directly to Aadhaar or SCS served as nodal points in their topical trajectories. In China, the U.S. subprime crisis became a part of the American Model topic and made it prominent in media coverage of SCS in 2007-08. Similarly, Alibaba's listing on foreign stock exchanges coincided with Market Regulation becoming a large topic in 2014. In India, the expected launch of the Direct Benefit Transfer (DBT) scheme in 2013 drew the media's attention to issues like food subsidies and payment transfers all through 2012 and made Public Welfare the dominant topic that year. And in 2014, the national election finally managed to get privacy issues related to Aadhaar on the media's agenda and turned Privacy Politics into a significant topic.

Durham (1998) has identified *positivism* as an ideology of journalism that underlies the framing process. He defines framing as "the mechanism for gathering and joining 'atomistic events' into news stories" (p. 114). From this perspective, frames connect disparate events to form social narratives and allow journalists to control how events are interpreted. The Indian and Chinese media's unwillingness to cover "non-events" and their co-optation of disparate domestic and international events into the framing of Aadhaar and SCS reflects this tendency. This is perhaps most obvious in how the U.S. sub-prime crisis was weaved into Chinese media coverage of the SCS and (mis)used to argue *for* the necessity of a credit-based system, when the crisis in fact showed how fragile such a system could be.

This points us toward yet another common ideological underpinning in the framing of Aadhaar and SCS — *neoliberalism*. News media in both India and China gave extensive coverage to the commercial benefits of these data projects and justified them on these very grounds. Previous studies have noted the Indian media's belief in "benign capitalism" (Shahin, 2017) and the Chinese media "wearing a socialist face with a capitalist body" (Lee, 2003, p. 18). Our analysis suggests that a neoliberal ideology also undergirds their framing of Big Data projects — in which public welfare is at best a byproduct of an economic system whose primary goal is to facilitate corporate profiteering.

Privacy and mass surveillance are conspicuously absent from the coverage in both countries. Aadhaar and SCS possess the potential to hand over the control of citizens' personal lives to governments and the private technology vendors they work in tandem with (Dixon, 2017; Hawkins, 2017). Yet in India, privacy becomes an important topic of coverage only in 2015. Even then, the coverage is not substantive. News reports focus on mundane court proceedings of privacy litigations rather than the significant concerns raised by the litigants. The issue is also framed within the context of the national election. In China, meanwhile, privacy never grows into a noteworthy topic as the news media disregard it altogether. It means that in both these nations, common citizens who rely on the news media for most — if not all — of their knowledge about these Big Data projects are unlikely to be aware of how these projects could undermine their personal lives and make them susceptible to government and corporate control.

The invocation of the so-called "freedom of individual choice" to hold individuals responsible for their own privacy in their dealings with technology companies and government-run data projects is thus naïveté at best subterfuge at worst. Individual choice is a chimera in a politico-economic milieu in which the news media, the most significant data intermediaries shaping public understandings of data and technologies, are institutionally and ideologically bound to the same governments and corporations that profit from data projects and technology services (Yu, 2017; Sen, 2016). Similarly, the construction of human beings as "privacy pragmatists" (Kumaraguru & Cranor, 2005) who weigh their loss of privacy against the benefits of using digital technologies or participating in data projects is simplistic and reductionist (Elueze & Quan-Haase, forthcoming; Hoofnagle & Urban, 2014). Our study shows that such "pragmatism" isn't individualistic and doesn't exist in a political and ethical vacuum: it is produced by a social structure that privileges technology adoption and participation in projects like Aadhaar and SCS while concealing how they compromise privacy and leave citizens susceptible to surveillance. Without such awareness, it is no surprise that most citizens "choose" to participate in these Big Data projects (Srinivasan et al., 2018).

Our analysis thus exposes how a nexus of neoliberal governments, corporations, and the news media imposes exploitative Big Data projects in nations of the Global South while making citizens structurally incapable of resisting them — or even recognizing the need to resist them. This relationship corresponds with what Lukes (1974, p. 24) called the *third face of power,* which makes citizens willing to "accept their role in the existing order of things" and nullifies even the possibility of resistance. This social structure is reflected in and reinforced through the semantic structure of news coverage — specifically the news framing process that shapes what citizens expect or want from these projects.

The social and policy implications of our study are wide-ranging. Both bottoms-up and top-down efforts are needed to alter the nature of structural relations that normalize surveillance and violations of privacy — although such efforts will have to be aligned with the political, economic, and cultural particularities of each nation. First, data protection and digital rights activists need to recognize the importance of educating citizens about why privacy matters and how Big Data projects undermine it. In effect, they ought to view themselves as data intermediaries who make positive efforts to build public awareness about data practices — rather than rely on the news media to do so. Only then could the public begin to have a meaningful "choice" in such matters. Such activism would be relatively easier in India than China, where the government has always clamped down hard on grassroots attempts to make digital technologies public service-oriented (Chen & Reese, 2015).

Second, lawmakers who are serious about data privacy need to challenge the received wisdom that individuals are or can be what Baruh and Popescu called the "locus of privacy decision" (2017, p. 585). Privacy ought to be treated as an inalienable right rather than a market commodity that individuals can trade for data-oriented services or conveniences. Moreover, those who provide such services — governments and corporations — ought to be made responsible for violations of privacy rights. These could be the first steps toward changing the governmental and corporate tendency to view data

and technology services as instruments of surveillance and a means to expand their social control.

This study has been critical of the news media's role as data intermediaries and identified how institutional and ideological linkages with governments and corporations make the media complicit in the exploitation of citizens' personal data. At the same time, we acknowledge that news framing is a social process in which multiple stakeholders can play a role. Therefore, a third implication of our study is that the news media — as data intermediaries — can also serve as allies in both top-down and bottoms-up efforts to defend data privacy. This, however, would not happen on its own. Concerned lawmakers and civil society groups need to productively engage with journalists and news organizations — for instance, by creating "events" that can serve as news pegs for privacy-related media coverage or by serving as sources of information for articles related to Big Data projects.

One of the strengths of our study — its broad geographical and temporal scope — is also its weakness. While enabling us to find common patterns of media coverage about Big Data projects across two very different media and political systems over time, our research design also prohibits us from delving too deeply into the analysis of any particular moment or topic. Future research may therefore study the more interesting topics or time periods that have emerged from our analysis in greater detail. For instance, the association of privacy and politics in Indian media coverage in 2014 merits closer attention. So does the Chinese media's misappropriation of the sub-prime crisis in 2008 to make the case for SCS. In addition, as governments across the Global South continue to launch "national" data or digitization projects, we hope that future research would expand the scope of our analysis by testing the relevance of our findings in new contexts.

REFERENCES

Acquisti, A., Gross, R. and Stutzman. (2011). Faces of Facebook: Privacy in the age of augmented reality. Presented at BlackHat Conference, Las Vegas, August 4.

Andrejevic, M., & Gates, K. (2014). Big data surveillance: Introduction. *Surveillance & Society*, *12*(2), 185-196.

Angwin, J., Larson, J., Mattu, S. & Kirchner, L. (2016). Machine bias. *ProPublica*. Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Baruh, L., & Popescu, M. (2017). Big data analytics and the limits of privacy self-management. *New Media & Society*, *19*(4), 579-596.

boyd, d., & Crawford, K. (2012). Critical questions for Big Data. *Information, Communication & Society*, *15*(5), 662–679.

Brewer, P. R., & Gross, K. (2010). Studying the effects of issue framing on public opinion about policy issues: Does what we see depend on how we look. In D'Angelo, P. & Kuypers, J. (Eds.) *Doing news framing analysis: Empirical and theoretical perspectives* (pp. 159-186). New York: Routledge.

Bloomberg. (2017). India ID program wins World Bank praise despite 'Big Brother' fears. Retrieved from https://www.bloomberg.com/news/articles/2017-03-15/india-id-program-wins-world-bank-praise-amid-big-brother-fears

Carragee, K. M., & Roefs, W. (2004). The neglect of power in recent framing research. *Journal of Communication*, 54(2), 214-233.

Chadha, M. S., Broor, S., Gunasekaran, P., Potdar, V. A., Krishnan, A., Chawla-Sarkar, M., ... & Klimov, A. (2012). Multisite virological influenza surveillance in India: 2004–2008. *Influenza and other respiratory viruses*, 6(3), 196-203.

Chakravartty, P., & Roy, S. (2013). Media pluralism redux: Towards new frameworks of comparative media studies "beyond the West". *Political Communication*, 30(3), 349-370.

Chandrasekhar, R. (2015). A shaky Aadhaar. *The Indian Express*. Retrieved from http://indianexpress.com/article/opinion/columns/a-shaky-aadhaar/

Chen, W., & Reese, S. D. (2015). Introduction: A new agenda: Digital media and civic engagement in Networked China. In W. Chen & S. D. Reese (Eds.) *Networked China: Global Dynamics of Digital Media and Civic Engagement* (pp. 21-36). New York: Routledge.

China Copyright and Media (2014). *Planning Outline for the Construction of a Social Credit System (2014-2020)*. Retrieved from https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020/.

Crawford, K. (2016). Artificial intelligence's white guy problem. *New York Times*. Retrieved from http://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html

Dai, L., Zhu, J., Liang, J., Wang, Y. P., Wang, H., & Mao, M. (2011). Birth defects surveillance in China. *World Journal of Pediatrics*, 7(4), 302-310.

De Vreese, C. H. (2010). Framing the economy: Effects of journalistic news frames. In D'Angelo, P. & Kuypers, J. (Eds.) *Doing news framing analysis: Empirical and theoretical perspectives* (pp. 187-214). New York: Routledge.

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. Government arts funding. *Poetics*, 41, 570–606.

Dixon, P. (2017). A Failure to Do No Harm – India's Aadhaar biometric ID program and its inability to protect privacy in relation to measures in Europe and the U.S. *Health Technology, 7*, 539-567.

Durham, F. S. (1998). News frames as social narratives: TWA Flight 800. *Journal of Communication*, 48(4), 100-117.

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51-58.

Federal Trade Commission (FTC) (2014) Data Brokers: A Call for Transparency and Accountability. Available at: https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport. pdf

Fung, A. Y. (2007). Political economy of Hong Kong media: Producing a hegemonic voice. *Asian Journal of Communication*, 17(2), 159-171.

Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, 95(1), 1-37.

Garson, B. (1988). *The electronic sweatshop: How computers are transforming the office of the future into the factory of the past.* Simon & Schuster, New York

Gellman, B., & Poitras, L. (2013). US, British intelligence mining data from nine US internet companies in broad secret program. *The Washington Post*. June 6. Retrieved from www.washingtonpost.com/investigations/usintelligence-mining-data-from-nine-us-internet-companies-in-broad-secret-program/2013/06/06/3a0coda8-cebf-11e2-8845-d97occb04497_story.html

Guo, L., Vargo, C. J., Pan, Z., Ding, W., & Ishwar, P. (2016). Big social data analytics in journalism and mass communication: Comparing dictionary-based text analysis and unsupervised topic modeling. *Journalism & Mass Communication Quarterly*, 93(2), 332-359.

Hawkins, A. (2017). Chinese Citizens Want the Government to Rank Them. *Foreign Policy*. Retrieved from http://foreignpolicy.com/2017/05/24/chinese-citizens-want-the-government-to-rank-them/

Herman, E. S., & Chomsky, N. (1988). *Manufacturing consent: The political economy of the mass media*. New York, NY: Pantheon Books.

Hindustan Times. (2017). Jharkhand's Aadhaar breach: India needs a strong data protection law. Retrieved from http://www.hindustantimes.com/editorials/jharkhand-s-aadhaar-breach-india-needs-a-strong-data-protection-law/story-50f3gapEnnMWiy7c7vJ4YN.html

Hoofnagle, C., & Urban, J. (2014). Alan Westin's privacy homo economicus. *Wake Forest Law Review,* 49, 261–309.

Xinhua.net (June 22, 2014). How much is trust. Retrieved from http://finance.ifeng.com/a/20140622/12585704_0.shtml

Indian Express. (2014). Don't exclude those without Aadhaar, share data: SC. Retrieved from http://indianexpress.com/article/india/india-others/supreme-court-bars-sharing-of-uidai-info-denial-of-benefits-to-those-without-aadhaar/

Jacobi, C., Atteveldt, W., & Welbers, K. (2016). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, 4, 89-106.

Jacobs, A. (2009) Pathologies of Big Data. *ACM Queue*, 7(6), July.

Kumaraguru, P., & Cranor, L. F. (2005). Privacy indexes: a survey of Westin's studies. *Institute for Software Research International*.

Lee, Chin-Chuan. (2003). *Chinese Media, Global Contexts*. London: Routledge.

Liang, X. F., Li, L., Liu, D. W., Li, K. L., Wu, W. D., Zhu, B. P., ... & Yin, D. P. (2011). Safety of influenza A (H1N1) vaccine in postmarketing surveillance in China. *New England Journal of Medicine*, 364(7), 638-647.

Lukes, S. (1974). *Power: A radical view*. London: Macmillan.

Marwick, A. E. (2018). 7053| Privacy at the Margins| Understanding Privacy at the Margins—Introduction. *International Journal of Communication*, 12, 1157–1165.

McIntyre, L., Michael, K., & Albrecht, K. (2015). RFID: Helpful new technology or threat to privacy and civil liberties?. *IEEE Potentials*, *34*(5), 13-18.

Mneney, J., & Van Belle, J. P. (2016, January). Big Data capabilities and readiness of South African retail organisations. In *Cloud System and Big Data Engineering (Confluence), 2016 6th International Conference* (pp. 279-286). IEEE.

Nilekani, N. (2015). Why Supreme Court judgment on Aadhaar calls for an appeal. *The Indian Express*. Retrieved from http://indianexpress.com/article/opinion/columns/why-supreme-court-judgment-on-aadhaar-calls-for-an-appeal/

Ohm, P. (2010) Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review, 57*, 1701-1777.

Pan, Z. (2000). Spatial configuration in institutional change: A case of China's journalism reforms. *Journalism*, *1*(3), 253-281.

Pentland, A. S. (2014). *Social physics: How good ideas spread—the lessons from a new science*. New York, NY: Penguin Books.

Philo, G. (2014). *The Glasgow media group reader, Vol. II: Industry, economy, war and politics.* London: Routledge.

Quail, C., & Larabie, C. (2010). Net neutrality: Media discourses and public perception. *Global Media Journal, 3*(1), 31-50.

Reese, S. D. (2001). Prologue—Framing Public Life. In Reese, S. D., Gandy Jr, O. H., & Grant, A. E. (Eds.). *Framing public life: Perspectives on media and our understanding of the social world* (7-31). New York: Routledge.

Roy, C., & Kalra H. (August 12, 2011). The information technology rules, 2011, PRS Legislative Research, Center for Policy Research. Retrieved from http://www.prsindia.org/uploads/media/IT%20Rules/ IT%20Rules%20and%20Regulations%20Brief%202011.pdf.

Safire, W. (2002). You are a suspect. *The New York Times*, November 14.

Sawicki, D. S., & Craig, W. J. (1996). The democratization of data: Bridging the gap for community groups. *Journal of the American Planning Association, 62*(4), 512-523.

Schrock, A., & Shaffer, G. (2017). Data ideologies of an interested public: A study of grassroots open government data intermediaries. *Big Data & Society*, *4*(1), 1-10.

Sen, B. (2016). *Digital politics and culture in contemporary India: The making of an info-nation*. London: Routledge.

Shah, B., & Mathur, P. (2010). Surveillance of cardiovascular disease risk factors in India: the need & scope. *The Indian Journal of Medical Research, 132*(5), 634.

Shahin, S. (2016a). When scale meets depth: Integrating natural language processing and textual analysis for studying digital corpora. *Communication Methods and Measures*, *10*(1), 28-50.

Shahin, S. (2016b). Right to Be Forgotten: How national identity, political orientation, and capitalist ideology structured a trans-Atlantic debate on information access and control. *Journalism & Mass Communication Quarterly*, *93*(2), 360-382.

Shahin, S. (2017). Facing up to Facebook: how digital activism, independent regulation, and mass media foiled a neoliberal threat to net neutrality. *Information, Communication & Society*. Published ahead of print. doi: 10.1080/1369118X.2017.1340494

Shahin, S., Zheng, P., Sturm, H. A., & Fadnis, D. (2016). Protesting the paradigm: A comparative study of news coverage of protests in Brazil, China, and India. *The International Journal of Press/Politics*, 21(2), 143-164.

Srinivasan, J., Bailur, S., Schoemaker, E., & Seshagiri, S. (2018). Privacy at the Margins| The Poverty of Privacy: Understanding Privacy Trade-Offs From Identity Infrastructure Users in India. *International Journal of Communication*, 12, 1228–1247.

Stanaland, A. J., & Lwin, M. O. (2013). Online privacy practices: Advances in China. *Journal of International Business Research*, *12*(2), 33.

Stoycheff, E. (2016). Under surveillance: Examining Facebook's spiral of silence effects in the wake of NSA Internet monitoring. *Journalism & Mass Communication Quarterly*, *93*(2), 296-311.

Sun, X., Wang, N., Li, D., Zheng, X., Qu, S., Wang, L., ... & Wang, L. (2007). The development of HIV/AIDS surveillance in China. *Aids*, *21*, S33-S38.

Tehelka. (2009). Falling between the bar codes. Retrieved from http://www.tehelka.com/2009/08/falling-between-the-barcodes/?singlepage=1

Thussu, D. K. (2007). The Murdochization of news? The case of Star TV in India. *Media, Culture & Society*, 29(4), 593-611.

UIDAI. (u.d.). About Aadhaar. Retrieved from https://uidai.gov.in/your-aadhaar/about-aadhaar.html

Ward, J. S., & Barker, A. (2013). Undefined by data: a survey of big data definitions. *arXiv preprint arXiv:1309.5821.*

White House (2012) Consumer data privacy in a networked world: A framework for protecting privacy and promoting innovation in the global digital economy. Available at: http://www. whitehouse.gov/sites/default/files/privacy-final.pdf

Wired. (2005). Why the world is flat. Retrieved from https://www.wired.com/2005/05/friedman-2/.

Wired. (2017). Big data meets Big Brother as China moves to rate its citizens. Retrieved from http://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion.

Yu, H. (2017). *Networking China: The Digital Transformation of the Chinese Economy*. University of Illinois Press.

Zhou, H (May 7, 2009). The urgency of building up financial credit system from: A lesson from US financial crisis. Retrieved form http://finance.people.com.cn/GB/67543/9260040.html