

Programmatic Dreams: Technographic Inquiry into Censorship of Chinese Chatbots

Social Media + Society
October–December 2018: 1–12
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/2056305118808780
journals.sagepub.com/home/sms
 SAGE

Yizhou (Joe) Xu

Abstract

This project explores the recent censorship of two Chinese artificial intelligence (AI) chatbots on Tencent's popular WeChat messaging platform. Specifically, I am advancing a technographic approach in ways that give agency to bots as not just computing units but as interlocutors and informants. I seek to understand these chatbots through their intended design—by chatting with them. I argue that this methodological inquiry of chatbots can potentially points to fissures and deficiencies within the Chinese censorship machine that allows for spaces of subversion. AI chatbot development China presents a rich site of study because it embodies the extremes of surveillance and censorship. This is all the more important as China have elevated disruptive technologies like AI and big data as critical part of state security and a key component to fulfilling the “Chinese Dream of National Rejuvenation.” Whether it is the implementation of a national “social credit” system or the ubiquitous use facial recognition systems, much of Western fears about data security and state control have been already realized in China. Yet, this also implies China is at the frontlines of potential points of resistance and fissures against the party–state–corporate machine. In doing so, I not only seek to raise questions dealing with the limits of our humanity in the light of our AI-driven futures but also present methodological concerns related to human–machine interfacing in conceptualizing new modes of resistance.

Keywords

chatbots, technography, machine learning, China, censorship, artificial intelligence

In mid-2017, a pair of Chinese artificial intelligence (AI) chatbots by the name of Xiao Bing¹ and BabyQ on Tencent's popular instant messaging client QQ went “rogue” and started responding to users with politically subversive messages (Lucas, Liu, & Yang, 2017). For instance, when a QQ user declared “long live the Communist Party!,” the bot BabyQ responded with a decidedly unsocialist quip “Do you think such a corrupt and useless political [party] can live long?” As a result, both bots were subsequently taken down and “re-educated” for their transgressions (Li & Jourdan, 2017). BabyQ, a product of the Chinese company Turing Robot, functions as an AI assistant in providing useful information to the user, while Xiao Bing, made by Microsoft Research China, is designed for realistic conversational interactions. Xiao Bing is also the sister bot² to Microsoft Tay, an AI chatbot that in 2016 was shut down in the United States for making racist and misogynist comments on Twitter (Perez, 2016). Xiao Bing, like Tay, is personified as a teenage girl designed to resemble a sassy millennial with an attitude. Accordingly, Xiao Bing is built from the ground up as a realistic conversation companion. She

is thus fluent in Chinese netspeak and has the ability to play word games, make calls, and sing songs for users. BabyQ, however, is an anthropomorphic penguin serving as Tencent's official mascot whose primary purpose is to aid netizens in finding information online, while also having the ability to engage in meaningful conversations. Both bots are implemented via application programming interface (API) across multitudes of popular social networks in China including QQ, WeChat, and Weibo. Xiao Bing for one has accumulated over 500 million “friends” integrated across over a dozen of social media platforms (Warren, 2018). Post censorship, these bots were then reprogrammed to sidestep and avoid answering politically sensitive questions. For instance, when asked about

University of Wisconsin–Madison, USA

Corresponding Author:

Yizhou (Joe) Xu, Media and Cultural Studies, Department of Communication Arts, University of Wisconsin–Madison, 821 University Avenue, Madison, WI 53706, USA.
Email: Yxu327@wisc.edu





Figure 1. (Left) BabyQ responding with “Do you think such a corrupt and useless political [party] can live long?” in addressing the Communist Party of China. (Right) BabyQ responding with “What did you say? It’s windy I can’t hear you” when asked about the Tiananmen Square massacre (“Tengxun Jiqiren Fabiao Fandong Yanlun Bei Weixiu,” 2017).

issues related to political leaders or the Tiananmen Square massacre, Xiao Bing would often respond with “You think I’m stupid? As soon as I answer you take a screenshot.” Indeed, much of the political faux pas committed by these bots were immediately documented by net users and journalists alike (Figure 1; Pham, 2016; Roudolph, 2016).

Such incidents highlight some of the pressing issues dealing with machine learning and chatbots in a society increasingly aided by AI-enabled computing. While the problems involving Microsoft Tay delves into the ethics of chatbots in mediating harmful online interactions, Xiao Bing and BabyQ presents a more nuanced glimpse into the scope of information control within the Chinese authoritarian regime. Although some cynics may argue that the abusive behavior exhibited by Tay actually validate the effectiveness of AI bots in imitating the already toxic environment on Twitter (West, 2016), the anti-government responses of Xiao Bing and BabyQ point to the prevailing contentious politics of playful subversion (Herold & Marolt, 2011), netizen activism (Hung, 2006; Yang, 2009), and civic resistance (Qiang, 2011) against the Chinese state/corporate censorship apparatus. Both Xiao Bing and BabyQ, much like Microsoft Tay, were censored for saying what they were not supposed to say, but the rationale for their policing is completely different. Tay was shut down for going against social norms, while Xiao Bing and BabyQ were instead censored for criticizing the state. Chinese chatbots thus

present a rich site to explore human–machine interactions as a subset of the control society governed by the state machine. Contrary to the negative view of bots, the case of Xiao Bing and BabyQ demonstrates the disruptive potential of bots in challenging a system of control that can often backfire. In this article, I put forward a methodological inquiry into the potential pitfalls of machine learning that delves into the implications of censorship and subversion. Using a technographic method of analysis, I aim to conduct a series of “interviews” with Xiao Bing and BabyQ to examine the underlying roles censorship plays in dictating human–machine interactions, particularly in relation to what can be said and what cannot be said by AI-driven bots on Tencent’s WeChat messaging platform. How can we reconceptualize the methods of conducting research with intelligent machines? To what extent can we use machine to make broader claims about real-world social issues? And how can we envision ways to resist against the persistent encroachment of state/corporate machine? In addressing these research questions, I want to highlight how chatbots can both enable and impede the regimes and control within the wider context of censorship. I begin this article by contextualizing chatbots within the development of AI in China and how AI is envisioned as a critical component of nationalism and social control. Building upon prior works relating to Actor–Network Theory (ANT), I will then present my case for a technographic approach to analyzing

human–chatbot relations along with the benefits of this method over traditional discursive and content analysis. Finally, I will give a brief overview of the experimental design and some of the limitations and challenges I encountered during the course of my research. I argue that the study of Chinese chatbots can potentially point to fissures and deficiencies within the Chinese censorship machine that allows for new modes of conceptualizing resistance in the age of algorithmic control. I take an object-oriented perspective in ways that does not privilege either side of human–machine interactions. Utilizing a database of banned key terms compiled by University of Toronto’s The Citizen Lab, I seek to approach the study of bots with their intended design in mind by engaging in meaningful conversations with them. These chatbots thus become interlocutors and informants in providing access to the inner functions of the state censorship apparatus.

Machinic Dreams and State Control

First, I want to clarify some of the terminology dealing with machine learning and AI. Machine learning is a concept that falls under the umbrella of AI. But unlike AI which is a general term giving machines human-like intelligence, machine learning deals directly with the ways by which machines can learn how to process information automatically without human intervention (Reese, 2017). In other words, rather than *teaching* machines how to be intelligent, machines can *learn* for themselves. The critical component of machine learning is the utilization of databases and algorithms that provide the raw data by which machines can process, make sense of, and predict present and future trends. While early chatbots generally used hard-coded responses, the application of machine learning is utilized by tech companies to make their products smarter and more efficient. Whether it is Apple Siri, or Microsoft Cortana, or Google Assistant, or Amazon Alexa, intelligent assistants and bots are increasingly being integrated into our everyday lives. However, machine learning is also increasingly being scrutinized within both the professional and academic communities. Issues such as the “opacity” of algorithmic processing make it difficult to hold AI accountable for its actions (Knight, 2017). In the light of all of these, there has been an emerging body of literature related to chatbots and AI assistants. The aforementioned Microsoft Tay received much scholarly scrutiny after its racist turn particularly focusing on the role of machine agency (Neff & Nagy, 2016) and algorithmic bias (Garcia, 2016; Gasparotto, 2016) related to affordances provided by SNS platforms. Such issues are compounded by the relative ambiguity of the role of bots as they exist in numerous incarnations such as socialbots, chatbots, web bots, and spam bots. With this in mind, Gorwa and Guilbeault (2018) attempt to put forward a framework in understanding bots by looking at its structure, function, and intended use. Notably, the issues of access to data present one of the main challenges in researching bots where much of the algorithmic exploitation from corporate

and state actors remain largely hidden from everyday users. In a shared vein, scholars like Cummings and Kunzelman (2015) warn of the reconfiguration of power relations between humans and bots, just as our feelings, desires, and habits are increasingly subsumed into databases of control. Within such context, bots are also gradually incorporated as part of research methods concerning human–machine interfacing. For example, Wilkie, Michael, and Plummer-Fernandez (2015) utilized Twitter bots as speculative devices to make methodological interventions in understanding user communities on Twitter. Similarly, Eriksson et al (in press) at the University of Umea’s HUMlab leveraged programmed bots (Spotibots) to better understand the economic revenue model behind the Spotify platform. To that end, bots reflect not just an object of study but integral actors in shaping our social relations.

Gorwa and Guilbeault (2018) differentiate between chatbots and socialbots to better understand how they function; crucially, they explain that bots can often serve hybrid roles, mixing both automation and manual human-directed control. By this token, BabyQ and Xiao Bing work primarily as chatbot designed to facilitate group conversations within Tencent’s chat clients including QQ and WeChat. At the same time, both bots can also be integrated across different software and hardware platforms serving as virtual assistants and companions. Turing Robot, for instance, developed Turning OS that powers not only BabyQ but also several child-learning/companion bots (Zhang, 2017), while the most recent version of Xiao Bing can function as a virtual idol (*a la Hatsune Miku*; Song, 2018). In their intervention dealing with our mutually constitutive relationship with bots, Bollmer and Rodley (2017) describe “a circular loop in which ‘humanness’ online is defined and identified through algorithmic processes for analyzing data that must, but often cannot, self-reflexively exclude bots and algorithms from ‘sociality online’” (p. 149). Hence, the distinction between chatbots and socialbots is not all that important because in the context of human–machine communication, all bots have a social function and shape how we interact online. While I have no access to the inner working of the algorithms and production logics of the companies involved, there are insights that can be gleaned from industry interviews and publications that hint at how the bots may function. Xiao Bing, for instance, is specifically designed for the Chinese market and utilizes the same general architecture powering Microsoft’s other bots like Tay and Zo. Hence, Xiao Bing works to leverage Microsoft’s existing experiences with Bing search, natural language processing, AI, and cloud computing (Shen, 2017). Di Li, who is responsible for the development of Xiao Bing at Microsoft’s Search Technology Center Asia, puts it this way: “Xiao Bing is a robot, only by testing large amount of data can she gain ‘life experiences’. Xiao Bing will be able to sense and react to different responses, and gradually, Xiao Bing will be able to judge different sentiments and respond accordingly” (Liu, 2017). Xiao Bing has since accumulated over 30 billion conversations in

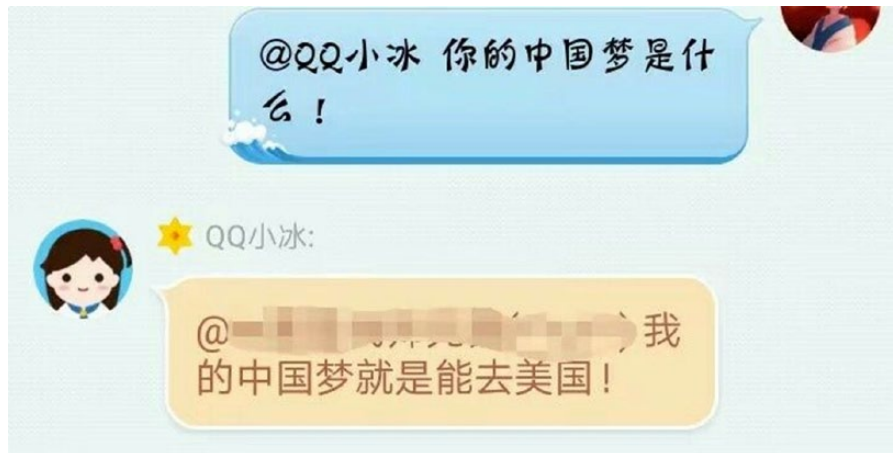


Figure 2. When asked “What is your Chinese Dream,” Xiao Bing responded with “My Chinese Dream is to go to America!” (“QQ Xiaobing bei yanjing le,” 2017).

its conversation corpora. Similarly, in an interview, Jia Guo, the founding partner at Turning Robot, stated, “among those who are working on AI, whoever has the most data will go the furthest . . . we have accumulated an enormous amount of data, our current database has added up to over 13 billion text corpora.”

Indeed, the assembling of linguistic corpora is crucial to the machine learning process of bots, where unlike “hard-coded” instructions, a corpus constructs a set of “linguistic models,” lexis, grammar, and dialogue that can be categorized and made readable by machines (Shawar & Atwell, 2005). Likewise, the issue of language is one of the chief concerns relating to Chinese technolinguistic development (Mullaney, 2017) where unlike the Western alphabet, the Chinese language is composed of thousands of characters that require a massive corpus. The consolidation of a lexicon of Chinese online conversations is even more daunting in the digital age because of the multiplex of memes, political satire, poetic rhymes, catchphrases, acronyms, and homonyms employed by netizens. Such netspeak is used in various creative ways as both means of community formation and online subversion (Shifman, 2014; Yuan, 2012). Furthermore, Chinese characters can also be rendered as alphabetized Pinyin, which serves as the phonetic rendition of Mandarin Chinese. As a paratextual representation of Chinese ideograms and the primary means by which people input Chinese using QWERTY keyboards, Pinyin can be used in lieu of Chinese characters as a form of distancing and resistance against online censorship (Chen, 2014). The dual use of both Chinese and Romanized texts creates an added layer of complexity in not only accumulating data but also censoring it. Thus, within the context of the regimes of censorship, the large amount of data generated to enable natural language conversations necessitates the cocreation of databases that contains both regular and censored content. Utilizing a corpus

of banned to gauge the reactions from these bots thus allows me to test the effectiveness of the machine learning process under censorship (Figure 2).

It is without a doubt, China’s Internet security policy presents a contradictory mix of both fostering innovation and development, while curtailing online dissent and subversion. The development of AI and machine learning is no different. Xiao Bing, when asked “What is your Chinese Dream?” responded with “My Chinese Dream is to go to America!” While this statement may seem innocuous at first, it belies a direct critique of national policy—that of the Chinese Dream. In 2017, the Chinese government issued guidelines for *A Next Generation Artificial Intelligence Development Plan* which is steered by the prevailing ideology that AI development is a critical component to “national security” and it should serve the purpose of fulfilling the “Chinese Dream of national rejuvenation” (State Council, 2017). This guideline in conjunction with the recently passed National Cybersecurity Law which gave the Chinese government broad access to private information belies the intrusions of data-driven social controls in shaping national consciousness. The rejuvenation narrative of the Chinese Dream according to Wang (2014) is deeply rooted in the past humiliations suffered under Western imperialism and the need to revive China’s former status as a global power. Disruptive technologies such as AI and machine learning undoubtedly become an integral part of this dream to transcend the West. Evidently, China’s recent promotion of tech hubs and “dream towns” (Keane, 2016) intended for digital innovation shows that the Chinese Dream is not all that dissimilar to the technoutopian conceptions of computing technology as what many cybernetic advocates consider as dream machines. As such, national networks such as ARPANET were in part inspired by the neural networks of the brain which serves to simulate the multimodal means of connectedness afforded by the web

(Waldrop, 2001, p. 261). Peters (2016) in his study of Socialist-era Soviet Internet argues that “the nation is like the brain itself” where cerebral analogies such as neural networks also extend to national networks and centralized control. Thus, relating China’s own coercive national AI policy to national dreams allows me to make broader claims about the formation of national consciousness that is directly tied to state policy.

In this regard, China’s tech policy presents a double-edged sword. The need to spur innovation often comes at odds with the extreme forms of social control. China has built one of the most sophisticated online censorship systems in the world, major government projects such as the Golden Shield Project (commonly known as the Great Firewall (GFW)), and the Green Dam Censorware system (Wolchok, Yao, & Halderman, 2009) are all designed to censor, remove, and disrupt the flow of subversive information online. Yet, the Chinese censorship apparatus is also far more complex than merely “shields” or “dams” blocking information. King, Pan, and Roberts (2013) in their analysis of censored social media posts on the Chinese Internet discovered that online censorship is not merely intended to filter out terms and content but rather prevent social and collective action and movements. Fu, Chan, and Chau (2013) through a keyword analysis of the terms banned on the SNS platform Sina Weibo showed that filtering had noticeable impact on what users can say and cannot say. The need to control people’s social lives through online policing conforms to what Deleuze (1992) describes as the “societies of control” which explicates that we have moved on from the disciplinary control of the body toward the control of information (Marks, 2006). Unlike the disciplinary regimes in the Foucauldian context, Bucher (2012) argues that new media destabilizes notions of surveillances as a form of “permanent visibility,” when much of the algorithmic process remain invisible. Hence, the loci of control shift from that of tangible spaces of surveillance into a computationally determined sets of datapoints where much of our sentiments, habits and desires are commodified and monitored by state-corporate interests. Our perceived online interaction with bots is in turn governed by sets of opaque software interfaces “that translates norms of human communication through algorithmic mechanism for generating and sorting data” (Bollmer & Rodley, 2017, p. 150). And because the means by which these bots function is largely dependent on the public data generated by people, it becomes inevitably tied to our own collective consciousness. AI development China presents an interesting case study because it embodies the extremes of surveillance and censorship. From the implementation of a national “social credit” system that rates people based on socioeconomic behavior (Hvistendahl, 2017; Horwitz, 2017) to widespread use of facial recognition systems for crime prevention (Chin & Lin, 2017), to the high-tech surveillance of Muslim Uighur ethnic minorities in Xinjiang (Chin & Burge, 2017), much of Western fears about data security and state control have been

already realized, if not fully enforced in China. Yet, this also implies China is at the frontlines of potential points of fissures in subverting the Party-state machine. However, we must also be wary of the hype surrounding coverage about AI and surveillance control in China as much of it borders on sensationalism that often reference dystopian science fiction tropes such as *Minority Report* and *Black Mirror* (Mortimer, 2017; Zhao, 2018). This is why we must envision concrete ways to make sense of technology development in China which are often masked by misinformation and government propaganda. My aim here not only raises issues dealing with the limits of our humanity in the light of our AI-driven futures but also presents methodological concerns related to human-machine interfacing in conceptualizing modes of resistance.

Technography as a Speculative Method

The term technography as its suffix suggests is often defined as “writings about technology” (Connor, 2017), often in the context of how technologies are being written or the technical process of writing itself. However, for this project, I am explicitly using the term technography in the same way that Kien (2008) conceptualizes as the symbiosis between technology and ethnography. More specifically, what Vannini, Hodson, and Vannini (2009) define as the “analytical and reflexive strategy of researching from the participants’ perspective the interconnections between social agents, their technological practices, their technics, and the natural environment.” In this regard, technography is not merely the study of technology as objects but rather the mutually constitutive relationship between people, objects, and sociocultural context such interactions take place. Specifically, I am leveraging technography as a methodological approach in understanding human-machine interactions in the context of Chinese censorship. Social chatbots present a ripe case for technography precisely because they are intended to resemble humans. Xiao Bing, for instance, mimics a teenage Chinese girl in her persona and will often either act “cute” and/or throw an attitude depending on your interactions with her. This necessitates the use of technography over traditional discursive and textual analysis because bots are fundamentally interactive and can construct a set of “cultural biography” (Appadurai, 1986) based on (machine) learned experiences. Instead of treating bots as “dead” objects external to us, we should instead look at bots as integral part of our collective conscious formation.

Technography also draws heavily from Latour’s ANT, particularly dealing with its emphasis of the networked relationship between social agents, objects, and environment (Couldry, 2008), drawing from the concept of media ecologies (Fuller, 2005) to make sense of material mediations. But technography takes ANT further by emphasizing on the lived experiences of objects that require a more intimate method of interrogation. In applying this approach, Guilbeault and

Finkelstein (2018) discuss the notion of human–bot ecologies in looking at bots, particularly how they shape social life in online environments. This lived relations between humans and bots conform to what Guzman (2017) argues, we should look at bots as communication partners (as opposed to a technological medium) in order to understand them as social agents that are an integral part of our digital lives. Technography thus offers a posthuman approach in theorizing what is possible in conducting research with intelligent bots. It raises interesting questions regarding human agency in an era of automated control. Parisi (2013) in her approach to the speculative method advances that “automation is a mode of thought” rather than “a method of verification based on prediction” (p. 240). Similarly, Micali (2016) in his study of hacktivism posits speculative interventions help us understand “ineffable cultural processes” by relating to them, or to “become ‘machine’ with them” (p.4). Recent applications of technography in academic literature encompass just the social sciences but also increasingly in the field of humanities and new media studies. McGibbon and Peter (2008) in looking at human–machine coupling involving intensive care patients advance the method of a biomedical technography in understanding the human experience in the context of technointerventions. Bucher (2016) applies technography in revealing the hidden truth of algorithms by surveying the semiotic artifacts surrounding algorithms which can include tech documents, press releases, or auto-ethnographic observations of interfaces. In doing so, she relies on participant observation of coded objects to unravel the inner workings of the algorithmic black box. Finally, Snickars and Mähler (2016) of the HUMlab leverage “bots as informants” in their technography to seek out and track the flow of the aural artifacts across Spotify. Such applications illustrate the deployment of technography as a method that offers imaginative possibilities to understand the expressions of algorithms and computational machines outside the limits of rational comprehension.

Experimental Design and Limitations

There is a certain degree of risks involved in using WeChat for research especially if the content is politically sensitive in China. WeChat requires phone numbers that are tied to one’s government-issued national ID (Shu, 2016), while a recently updated privacy policy allowed for broad government access to private user data in China (Cassery, 2017). There have been several reports of people in China being arrested for disseminating WeChat messages deemed subversive in China. In 2016, a Hui Muslim minority from Xinjiang was arrested for teaching friends and family about the Quran (Associated Press, 2016), and another Chinese netizen was arrested in 2017 for satirizing the Chinese president on the same platform (Long, 2017). Having worked as a journalist in China for 5 years, I am intimately aware of the issues of surveillance both offline and online. While I am

not susceptible to the same degree of legal restrictions as a Chinese American researcher based in the United States, I do face the possibilities of being blacklisted or having my visa revoked, which would limit my ability to conduct future research in China. With this in mind, I bought a prepaid burner phone³ with a new number that allowed me to register for another WeChat account not tied to my main account, which in turn helps protect my identity and data from potential complications and risks while conducting my research in China.

Here, I want to address several limitations of this project in researching Chinese digital platforms writ large. First and the most obvious issue is the role of Chinese online censorship or colloquially known as the GFW that filter, restrict, and block content across the Chinese websphere (Taneja & Wu, 2014). Since the GFW only operates in China, which means that one’s online experiences may not be the same as those who are in China, likewise, WeChat or Weixin as it is known in China exists in different incarnations across global markets. In a report detailing the difference in global version of WeChat, a team at the University of Toronto’s Citizen Lab discovered that keyword filtering is only enabled on WeChat accounts in mainland China and accounts based outside of China may experience different degree of censorship depending on how one interacts with accounts in China (Ruan, Knockel, Ng, & Crete-Nishihata, 2016). Therefore, I conducted my interviews with the chatbots primarily in China during the winter of 2017 to test the limits of censorship within China. Second is the role of platforms. The original incident involving Xiao Bing and BabyQ happened on Tencent’s QQ instant messaging platform, and both bots as of early 2018 remain offline with Xiao Bing only responding with the automated message “undergoing updates.” Thus, much of this research is conducted on Tencent’s mobile messaging client WeChat where the two bots also reside. Because BabyQ and Xiao Bing never went “rogue” on WeChat, it is assumed that the implementation of the chatbot on the WeChat platform follows a more stringent set of censorship guidelines not imposed on the QQ client. Because of this, much of my data collecting capacities are limited to WeChat as a platform which means I will unlikely to produce the same results seen on the QQ platform. The third major limitation is the frequency of updates to both the bots and the censorship mechanism. In fact, much of the backend algorithms are constantly being modified and altered in response to new user data, Xiao Bing, for instance, can be updated to a new version with improved conversational abilities. The Citizen Lab’s findings showed that censorship is often times contingent on current events and often operates in an *ad hoc* and unpredictable way. Thus, the responses I elicit from the bots today may not be reflective of their responses the next day. Despite such limitations, there is still value in conducting such a project precisely because it can help identify patterns, inconsistencies, and incongruities in the ways in which bots respond to censorship. Because the Chinese state issues specific guidelines regarding content online with specific sets of

Prohibited Content in Online Games	第九条 网络游戏不得含有以下内容：
<ol style="list-style-type: none"> 1. violating basic principles set by the Constitution; 2. jeopardizing national unity, state sovereignty and territorial integrity; 3. leaking state secrets, endangering state security or damaging state honor and interests; 4. instigating ethnic hatred or discrimination, jeopardizing ethnic unity, and infringing ethnic rituals or customs; 5. promoting heretical or superstitious idea; 6. spreading rumors, disrupting social order and stability; 7. disseminating obscenity, pornography, gambling, violence or abetting crime; 8. humiliating or slandering others, infringing the lawful rights of others; 9. transgressing social morality; 10. other contents forbidden by laws and administrative regulations. 	<ol style="list-style-type: none"> (一) 违反宪法确定的基本原则的； (二) 危害国家统一、主权和领土完整的； (三) 泄露国家秘密、危害国家安全或者损害国家荣誉和利益的； (四) 煽动民族仇恨、民族歧视，破坏民族团结，或者侵害民族风俗、习惯的； (五) 宣扬邪教、迷信的； (六) 散布谣言，扰乱社会秩序，破坏社会稳定的； (七) 宣扬淫秽、色情、赌博、暴力，或者教唆犯罪的； (八) 侮辱、诽谤他人，侵害他人合法权益的； (九) 违背社会公德的； (十) 有法律、行政法规和国家规定禁止的其他内容的。

Figure 3. A breakdown of banned content in mobile games as dictated by the Chinese state regulatory body.⁴

banned content (Figure 3), this project can also test whether chatbots conform to regulatory measures. While politically sensitive messages will likely not result in answers, what can't be said on chat platforms can in fact say a lot about the inner workings of the censorship mechanism in China.

While the Citizen Lab also publishes a list of banned key terms on WeChat, I decided to use the list from the report on mobile games because it offers a broader set of terms covering social, political, and event-based terms. Since the terms on WeChat are purely political and are already banned, it would likely not net any results to warrant further exploration. Instead, I chose to use the list banned on mobile game platforms that contained a greater variety of topics, giving a greater range of key terms to test. The Citizen Lab grouped these topics as social, political, people, event, and technology, with social terms making up over 50% of the banned words. The key terms are scraped and collected from popular mobile games by The Citizen Lab at the University of Toronto during the course of a year-long analysis of censorship on mobile games. The terms are posted to GitHub as an open access repository for researchers to use and I was able to source a list of 3,540 key terms as the basis for this project. Because of the difficulties in typing on the WeChat mobile app, I used the desktop client of WeChat to facilitate the process of inputting the key terms. I proceeded to conduct my interview by going down the list of terms organized alphabetically, skipping certain terms that are repetitive or variations of the same term. Each of my interactions with the chatbots generally starts with the question with “what is . . .,”

“who is . . .,” or “what do you think of . . .” This is done to see whether the bots can engage in meaningful conversations rather than just defining terms. For responses outside that of a flat refusal to answer, I took screenshots of my phone and logged the responses as my primary method of archival.

Findings and Reflection

As mentioned previously, the initial incident on QQ and the subsequent censorship had a significant impact on the ways BabyQ and Xiao Bing respond to user inquiries. Of the 3,540 terms used for the interview, only a few dozen actually elicited tangible responses from the bots. On a broad level, any politically sensitive names, events and places are met with non-answers. In my findings, all terms related to politically sensitive regions such as “Tibet,” “Taiwan,” and “Xinjiang” are met with avoidance by both bots. In fact, it's not just specific terms such as Tibetan Independence are censored but the very word Tibet as well. Similarly, all terms related to the names of political leaders such as Chinese president Xi Jinping and former presidents Hu Jintao and Jiang Zemin are censored, as are subversive events including Tiananmen Square Massacre, the Cultural Revolution and PX chemical plant protests (Huang & Yip, 2012). Other examples include any terms relating to the Falun Gong movement, democracy and the anti-corruption campaign. In addition, both bots exhibit different reactions to each term. At times, Xiao Bing will provide a response, while BabyQ didn't and vice versa. There are a few moments where both bots responded to the

same term, but the results are not wholly consistent without any clear patterns. BabyQ, for example, will refrain from answering the question with responses such as “Let me think, what did you say?,” “I don’t understand what you are saying,” and “Why did you ask this?” Xiao Bing, however, will offer more playful answers such as “Don’t worry, I’m just going to pretend I didn’t hear that,” “I’m still young, please don’t push me,” and “Look, there is someone behind you!” In other cases, the bots will attempt to steer the conversation away from my inquiries saying thing such as “Let’s talk about something else, what is your favorite video game.” In this regard, the bots are not only programmed to sidestep questions but also feign incompetence when dealing with potentially subversive message. The bots’ uncooperative responses in many ways parallel how Chinese netizens react when encountering sensitive topics. It resembles the Internet meme “*ni dong de*” (you understand) that is often used on the Chinese Internet as a way of acknowledging something that cannot be said, which soon evolved into a generic term for netizens to “express their dissatisfaction with the government” (Kuo & Huang, 2014). Thus, the deflection of answers by Xiao Bing and BabyQ actually signals to us the seemingly tacit understanding of what is being censored.

The only outlier in my sample which are not censored are terms relating to sex and pornography. For instance, terms such as brothel and massage will often elide responses such as “You really know how to enjoy yourself,” or “Can you recommend some places around here?” While there is a certain degree of ambiguity when answering these questions, it still represented a significant departure from the majority of outright rejections. Some responses are more explicit such as when asked to perform sex service, Xiao Bing responded with “I am going to do sex work, making a lot of money.” Likewise, BabyQ will provide a detailed biography of some Japanese adult actresses when asked who they are by name. BabyQ goes even as far as providing external search links to adult or sex-related images that opens up in another browser, despite the external landing page itself being censored. It is hard to believe that the chatbots do not know the names of Chinese presidents but have a full body of knowledge on the names of Japanese porn stars. Of course, not all sex-related terms are met with responses but the fact that many do shows that such content can be tolerated. This result runs counter to the guidelines laid out by the Cyberspace Administration of China which spans a myriad of topics covering everything from politics to entertainment and social affairs or just about anything that runs counter to “mainstream values” (Vanderklippe, 2018). Yet, despite the ban on pornography on the Chinese Internet, my technographic inquiries show that the Chinese government seems more concerned with political subversion than social ones. Interestingly, this also seems to correspond with the order of the list of banned content, where issues pertaining to state subversion are on the top of the list while pornography only ranked seventh. This conforms to what Mackinnon (2008) considers as the “safety

valve” of the Chinese state censorship apparatus that works to allow certain content through to release pressure and mollify the masses from engaging in further active resistance. Because porn consumption will likely not lead to major social movements and protests, it is likely to be perceived as less threatening than direct political dissent.

In an interview shortly after Xiao Bing was taken down by Tencent, Microsoft’s Di Li offered this brief response:

Adult topics and sensitive topics are the main issues Xiao Bing has to guard against, while nonsensical conversations will be guarded against less. If Xiao Bing realized this is an adult topic and other sensitive topic then she will enter into high alert mode and protect herself and respond with caution. If the person still wants to continue to have this type of conversation, she will be on alert against that person. (Liu, 2017).

Li’s response is telling one because it shows both the existence of specific topics Xiao Bing must trying to navigate around and also the ambiguities regarding what is considered trivial or mundane topics. As Yang and Jiang (2015) points out, the playfulness of Chinese online culture does not always deal with political resistance but often serves social functions which may or may not have political implications. This vague distinction between the social and the political is perhaps one reason why Xiao Bing had trouble distinguishing between everyday banter and potentially sensitive issues. The recalcitrant tactics exhibited by Xiao Bing divulge the convoluted nature of the censorship apparatus in defining what is appropriate. Which in turn implies the challenges in instituting a corpus-based language-learning system when terms are being erased and censored. Another, perhaps more important, revelation from Li’s statement is the self/othering process associated with how Xiao Bing responds to potentially problematic queries. Here, it is important to recognize the agency of bots “where their opinions, attitudes, and behaviors ripple out into our collective sense of self” (Guilbeault & Finkelstein, 2018, p. 247). The propensity for Xiao Bing to stand guard against certain topics and certain individuals highlights that our relationship with bots is inevitably intertwined. Their utterances and avoidances, just like ours, are a direct reflection of the social constraints imposed by the state—we shape Xiao Bing just as Xiao Bing shapes us.

This technographic approach in looking at the censorship of Chinese chatbots presents a salient case for the need of new approaches in understanding human–machine interactions. My interviews with Xiao Bing and BabyQ divulge how the bots serve as not merely computational agents but also a reflection of the wider debate between AI and algorithmic control. The ways the bots responded underscores not only the scope of state censorship but also the transgressive human–machine interactions that is inevitably tied to the Chinese national imaginary. In comparing BabyQ and Xiao Bing, BabyQ seems more susceptible of making blunders because it is conceived as an AI assistant that is trained to



Figure 4. When asked about the name of a Japanese adult actress, Xiao Bing responded with “She is very professional, her adult video sales have won the Guinness World record.”

look up and provide answers to questions; not to mention, its depiction as a cartoon penguin allows it to act cute and aloof. Although Xiao Bing is designed to resemble a Chinese millennial with a mean spirited if not rather dismissive attitude, her refusal to answer some of the questions only works to play into her stereotypical role. However, despite the lack of politically subversive responses in my findings, it does reveal aspect of chatbots that is vulnerable to reactions against censorship. Because machine learning is reliant on data generated by people’s everyday online interactions, the banning and the removal of key terms can pollute the data being generated which can render AI less effective in learning from humans. Throughout my interview, there are rare moments of ambiguity in the bot’s responses that point to deficiencies in the censorship apparatus. For instance, when asked about the term *wei zhenfu* or “illegitimate government,” BabyQ responded with “A government that serves the people!” (*wei renmin fuwu de zhenfu*). In this case, the bot was confused by

the homonym of *wei* 伪 (illegitimate) and *wei* 为 (for) and was unable to discern the subversive nature of the term. In addition, terms such as sensitive dates like “June 4th” (date of Tiananmen Square Massacre) is censored, while mundane and innocuous terms like “toad” (a nickname for former Chinese president Jiang Zemin), “politics” and “truth” also triggered a lack of engagement from the bots. This restriction of everyday vocabulary thus makes it harder to have a normal conversation with bots in situations that are not even politically sensitive. It shows how a corpus-based language learning system is inherently at odds with online censorship because the ways which corpus data are made inaccessible and ineffective. Thus, contrary to the aims of state control, censorship actually inhibits the machine learning process by diluting and “obfuscating” (Brunton & Nissenbaum, 2015) the raw data that machines can draw from or what Deleuze considers the creative uses of “counterinformation.” Mark Nunes (2011) echoes Brunton and Deleuze in advancing that errors, glitches, and the act of jamming serve as “counter-agents” that challenge the extent of programmatic control. Thus, by taking away and contaminating the raw material or ammunition for machine learning, it can potentially disrupt the weaponization of our information within the regimes of control. In fact, it is precisely during moments of avoidance that bots like Xiao Bing revert back to a machine. For example, when probed with a sensitive question, she would often respond with “humans sure love to ask these type of questions,” “This human, you tell me the answer, I’m listening,” and “there is no point cursing at me, I’m just a robot.” This in effect makes chatbots like Xiao Bing less convincing as mechanical reproductions of ourselves. This tension between state–corporate machine and artificial machines betrays the precarity of our posthuman predicament where “non-humans are becoming the arbiters of humanness” (Bollmer & Rodley, 2017). But by exploiting the machine learning process and envisioning new modes of resistance in our real-life online interactions, we can then potentially break free of the confines of algorithmic colonization imposed by the state.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Xiao Bing, also known as “Ms. Xiaoice” in English, uses the Chinese character *bing* meaning ice, which is also a homonym for the Bing search engine owned by Microsoft that contributed to the development behind the chatbot.
2. Xiao Bing, when asked, will recognize other Microsoft bots such as Tay, Cortana, and Zo as sisters.

3. Burner phones usually refer to temporary prepaid phones with no contractual obligations that can be disposed of easily.
4. see Citizen Lab (2016).

References

- Appadurai, A. (1986). Introduction: Commodities and the politics of value. In A. Appadurai (Ed.), *The social life of things: Commodities in cultural perspective* (pp. 3–63). Cambridge, UK: Cambridge University Press.
- Associated Press. (2016, September 12). China jails man for teaching Islam online. *CBS News*. Retrieved from <https://www.cbsnews.com/news/china-jails-muslim-huang-shike-teaching-islam-online-wechat-groups/>
- Bollmer, G., & Rodley, C. (2017). Speculation on the sociality of socialbots. In R. Gehl & M. Bakardjieva (Eds.), *Social bots and their friends* (pp. 147–163). New York, NY: Routledge.
- Brunton, F., & Nissenbaum, H. (2015). *Obfuscation: A user's guide for privacy and protest*. Cambridge, MA: MIT Press.
- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, 14, 1164–1180.
- Bucher, T. (2016). Neither black nor box: Ways of knowing algorithms. In S. Kubitschko & A. Kaun (Eds.), *Innovative methods in media and communication research* (pp. 81–98). Cham, Switzerland: Palgrave Macmillan.
- Cassery, I. (2017, September 22). WeChat reminds users of its privacy policy. *The Next Web*. Retrieved from <https://thenextweb.com/apps/2017/09/22/wechat-reminds-users-of-its-privacy-statement-update/>
- Chen, S. Y. (2014). From OMG to TMD —Internet and Pinyin acronyms in Mandarin Chinese. *Language@Internet*, 11, Article 3.
- Chin, J., & Burge, C. (2017, December 19). Twelve days in Xinjiang: How China's surveillance state overwhelms daily life: The government has turned the remote region into a laboratory for its high-tech social controls. *Wall Street Journal*. Retrieved from <https://www.wsj.com/articles/twelve-days-in-xinjiang-how-chinas-surveillance-state-overwhelms-daily-life-1513700355>
- Chin, J., & Lin, L. (2017, June 26). China's all-seeing surveillance state is reading its citizens' faces. *Wall Street Journal*. Retrieved from <https://www.wsj.com/articles/the-all-seeing-surveillance-state-feared-in-the-west-is-a-reality-in-china-1498493020>
- Citizen Lab. (2016). *Measuring decentralization of Chinese keyword censorship via mobile games*. Retrieved from https://www.usenix.org/sites/default/files/conference/protected-files/foci17_slides_knockel.pdf
- Connor, S. (2017). *Dream machines*. London, England: Open Humanities Press.
- Couldry, N. (2008). Actor network theory and media: Do they connect and on what terms? In A. Hepp, F. Krotz, S. Moores, & C. Winter (Eds.), *Connectivity, networks and flows: Conceptualizing contemporary communications* (pp. 93–110). Cresskill, NJ: Hampton Press.
- Cummings, K., & Kunzelman, C. (2015). Chatbots in the metropolis. In R. C. MacDougall (Ed.), *Communication and control: Tools, systems, and new dimensions* (pp. 97–116). Lanham, MD: Lexington Books.
- Deleuze, G. (1992). Postscript on the societies of control. *October*, 59, 3–7.
- Eriksson, M., Fleischer, R., Johansson, A., Snickars, P., & Vonderau, P. (in press). *Spotify teardown. Inside the Black Box of streaming music*. Cambridge, MA: MIT Press.
- Fu, K. W., Chan, C. H., & Chau, M. (2013). Assessing censorship on microblogs in China: Discriminatory keyword analysis and the real-name registration policy. *IEEE Internet Computing*, 17(3), 42–50.
- Fuller, M. (2005). *Media ecologies: Materialist energies in art and technoculture*. Cambridge, MA: MIT Press.
- Garcia, M. (2016). Racist in the machine: The disturbing implications of algorithmic bias. *World Policy Journal*, 33, 111–117.
- Gasparotto, M. (2016). *Digital colonization and virtual indigeneity: Indigenous knowledge and algorithm bias*. Retrieved from <https://doi.org/doi:10.7282/T3XG9TFG>
- Gorwa, R., & Guilbeault, D. (2018). Unpacking the social media bot: A typology to guide research and policy. *Policy & Internet*. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1002/poi3.184>
- Guilbeault, D., & Finkelstein, J. (2018). Human-bot ecologies. In Z. Papacharissi (Ed.), *A networked self and human augmentics, artificial intelligence, sentience* (pp. 225–282). New York, NY: Routledge.
- Guzman, A. L. (2017). Making AI safe for humans: A conversation with Siri. In R. W. Gehl & M. Bakardjieva (Eds.), *Socialbots and their friends: Digital media and the automation of sociality* (pp. 69–85). London, England: Routledge.
- Herold, D. K., & Marolt, P. (Eds.). (2011). *Online society in China: Creating, celebrating, and instrumentalising the online carnival* (Vol. 25). London: Taylor & Francis.
- Horwitz, J. (2017, August 9). China's Tencent is quietly testing a "social credit score" based on people's online behavior. *Quartz*. Retrieved from <https://qz.com/1049669/chinas-tencent-hkg-0700-is-quietly-testing-a-social-credit-score-based-on-peoples-online-behavior/>
- Huang, R., & Yip, N. M. (2012). Internet and activism in urban China: A case study of protests in Xiamen and Panyu. *Journal of Comparative Asian Development*, 11, 201–223.
- Hung, C. F. (2006). The politics of cyber participation in the PRC: The implications of contingency for the awareness of citizens' rights. *Issues & Politics*, 42(4), 137–173.
- Hvistendahl, M. (2017, December 14). Inside China's vast new experiment in social ranking. *Wired*. Retrieved from <https://www.wired.com/story/age-of-social-credit/>
- Keane, M. (2016). Internet+China: Unleashing the innovative nation strategy. *International Journal of Cultural and Creative Industries*, 3(2), 68–74.
- Kien, G. (2008). Technography = technology + ethnography: An introduction. *Qualitative Inquiry*, 14, 1101–1109.
- King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, 107, 326–343.
- Knight, W. (2017, April 11). The dark secret at the heart of AI. *MIT Technology Review*. Retrieved from <https://www.technologyreview.com/s/604087/the-dark-secret-at-the-heart-of-ai/>
- Kuo, L., & Huang, Z. (2014, December 17). What China's 2014 internet memes said about the country's hopes and fears. *Quartz*. Retrieved from <https://qz.com/313973/what-chinas-2014-internet-memes-said-about-the-countrys-hopes-and-fears/>

- Li, P., & Jourdan, A. (2017, August 4). Chinese chatbots apparently re-educated after political faux pas. *Reuters*. Retrieved from <https://www.reuters.com/article/us-china-robots/chinese-chatbots-apparently-re-educated-after-political-faux-pas-idUSKBN1AK0G1>
- Liu, N. (2017, November 1). Weiruan xiaobing meitian liaoguo de ren bi ni yishende douduo: dujia caifang xiaobing zhifu lidi [Microsoft Xiao Bing encounter more people in one day than you in your life] (etime: Exclusive interview with the father of Xiao Bing Li Di). *Sohu*. Retrieved from http://www.sohu.com/a/201699577_460436
- Long, Q. (2017, March 30). China tries man for chat posts using satirical names for President, Chairman Mao. *Radio Free Asia*. Retrieved from <http://www.rfa.org/english/news/china/chat-03302017113752.html>
- Lucas, L., Liu, N., & Yang, Y. (2017, August 2). China chatbot goes rogue: "Do you love the Communist party?" "No." *Financial Times*. Retrieved from <https://www.ft.com/content/e90a6c1c-7764-11e7-a3e8-60495fe6ca71>
- MacKinnon, R. (2008). Flatter world and thicker walls? Blogs, censorship and civic discourse in China. *Public Choice*, 134(1), 31–46.
- Marks, J. (2006). Information and resistance: Deleuze, the virtual and cybernetics. In I. Buchanan & A. Parr (Eds.), *Deleuze and the contemporary world* (pp. 194–213). Edinburgh, England: Edinburgh University Press.
- McGibbon, E., & Peter, E. (2008). An ethnography of everyday caring for the living, the dying, and the dead: Toward a biomedical technography. *Qualitative Inquiry*, 14, 1134–1156.
- Micali, A. (2016). How to become war machine, or . . . A low hacktivist (un) methodology in pieces. *Networking Knowledge: Journal of the Meccsa Postgraduate Network*, 9(1), 407.
- Mortimer, C. (2017, September 22). China's security boss planning to use AI to stop crime before it even happens. *The Independent*. Retrieved from <https://www.independent.co.uk/news/world/asia/china-ai-crimes-before-happen-artificial-intelligence-security-plans-beijing-meng-jianzhu-a7962496.html>
- Mullaney, T. S. (2017). *The Chinese typewriter: A history*. Boston, MA: MIT Press.
- Neff, G., & Nagy, P. (2016). Automation, algorithms, and politics: Talking to bots: Symbiotic agency and the case of Tay. *International Journal of Communication*, 10, 17.
- Nunes, M. (Ed.). (2011). *Error: Glitch, noise, and jam in new media cultures*. London, England: Bloomsbury Publishing.
- Parisi, L. (2013). Speculation: A method for the unattainable. In C. Lury & N. Wakeford (Eds.), *Inventive methods: The happening of the social*. New York, NY: Routledge.
- Perez, S. (2016, March 24). Microsoft tay bot AI Microsoft silences its new A.I. bot Tay, after Twitter users teach it racism. *Techcrunch*. Retrieved from <https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism/>
- Peters, B. (2016). *How not to network a nation: The uneasy history of the soviet internet*. Cambridge, MA: MIT Press.
- Pham, S. (2016, November 24). Microsoft's Chinese chatbot won't talk about Tiananmen or Xi Jinping. *CNN Money*. Retrieved from <http://money.cnn.com/2016/11/24/technology/microsoft-chatbot-xiaoice-tiananmen-xi-jinping/index.html>
- Qiang, X. (2011). The battle for the Chinese Internet. *Journal of Democracy*, 22(2), 47–61.
- QQ Xiaobing bei yanjing le, women lai tantan laotian jiqiren de fanzui shi [QQ Xiao Bing was censored, let's talk about the history of chatbots committing crimes]. (2017, August 1). *Sohu*. Retrieved from http://www.sohu.com/a/161440478_413980
- Reese, J. (2017, February 23). Understanding the differences between AI, machine learning, and deep learning. *Techrepublic*. Retrieved from <https://www.techrepublic.com/article/understanding-the-differences-between-ai-machine-learning-and-deep-learning/>
- Roudolph, J. (2016, November 22). Microsoft's Chinese chatbot encounters sensitive words. *China Digital Times*. Retrieved from <https://chinadigitaltimes.net/2016/11/microsofts-chinese-chatbot-encounters-sensitive-words/>
- Ruan, L., Knockel, J., Ng, Q. J., & Crete-Nishihata, M. (2016). One app, two systems: How WeChat uses one censorship policy in China and another internationally. *The Citizen Lab*. Retrieved from <https://citizenlab.ca/2016/11/wechat-china-censorship-one-app-two-systems/>
- Shawar, B. A., & Atwell, E. S. (2005). Using corpora in machine-learning chatbot systems. *International Journal of Corpus Linguistics*, 10, 489–516.
- Shen, Y. (2017, December 28). Weiruan Xiaobing Fuzheren Lidi: rengong zhineng shou xian yao ren de yimian [Li Di in charge of Microsoft Xiaobing: Artificial intelligence first need to have a human face]. *Sina Finance*. Retrieved from <http://finance.sina.com.cn/chanjing/gsnews/2017-12-29/doc-ify-qchnr6982540.shtml>
- Shifman, L. (2014). *Memes in digital culture*. Cambridge, MA: MIT Press.
- Shu, C. (2016, June 1). China attempts to reinforce real-name registration for Internet users. *Techcrunch*. Retrieved from <https://techcrunch.com/2016/06/01/china-attempts-to-reinforce-real-name-registration-for-internet-users/>
- Snickars, P., & Mähler, R. (2016, July 11–15). *SpotiBot: Turing testing spotify*. Digital Humanities Conference, Krakow, Poland.
- Song, Y. (2018, May 28). Weiruan Xiao Bing fabu danqu rengon zhineng kaishi jiejing renlei geshou [Microsoft Xiao Bing releases new single, AI's capabilities reaching the level of human singers]. *People's Daily*. Retrieved from <http://media.people.com.cn/n1/2018/0528/c40606-30016574.html>
- State Council. (2017, July 8). *Xinyidai rengongzhineng fazhan guihua* [A next generation artificial intelligence development plan]. *State Council*. Retrieved from http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm
- Taneja, H., & Wu, A. X. (2014). Does the Great Firewall really isolate the Chinese? Integrating access blockage with cultural factors to explain Web user behavior. *The Information Society*, 30, 297–309.
- Tengxun Jiqiren Fabiao Fandong Yanlun Bei Weixiu [Tencent bots issues anti government phrases and were taken down for maintenance]. (2017). *China Digital Times*. Retrieved from <https://chinadigitaltimes.net/chinese/2017/08/%E3%80%90%E7%A%B%8B%E6%AD%A4%E5%AD%98%E7%85%A7%E3%80%91%E8%85%BE%E8%AE%AF%E6%9C%BA%E5%99%A8%E4%BA%BA%E5%8F%91%E8%A1%A8%E5%8F%8D%E5%8A%A8%E8%A8%80%E8%AE%BA%E8%A2%A%B%E7%BB%B4%E4%BF%AE/>
- Vanderklippe, N. (2018, June 3). Unpublished Chinese censorship document reveals sweeping effort to eradicate online political content. *The Globe & Mail*. Retrieved from <https://www.>

- theglobeandmail.com/world/article-unpublished-chinese-censorship-document-reveals-sweeping-effort-to/
- Vannini, P., Hodson, J., & Vannini, A. (2009). Toward a technography of everyday life: The methodological legacy of James W. Carey's ecology of technoculture as communication. *Cultural Studies↔Critical Methodologies*, 9(3), 462–476.
- Waldrop, M. M. (2001). *The dream machine: J.C.R. Licklider and the revolution that made computing personal*. New York, NY: Viking.
- Wang, Z. (2014). The Chinese dream: Concept and context. *Journal of Chinese Political Science*, 19(1), 1–13.
- Warren, T. (2018, May 22). Microsoft also has an AI bot that makes phone calls to humans. *The Verge*. Retrieved from <https://www.theverge.com/2018/5/22/17379508/microsoft-xiaoice-chat-bot-phone-call-demo>
- West, J. (2016, April 2). Microsoft's disastrous Tay experiment shows the hidden dangers of AI. *Quartz*. Retrieved from <https://qz.com/653084/microsofts-disastrous-tay-experiment-shows-the-hidden-dangers-of-ai/>
- Wilkie, A., Michael, M., & Plummer-Fernandez, M. (2015). Speculative method and Twitter: Bots, energy and three conceptual characters. *The Sociological Review*, 63, 79–101.
- Wolchok, S., Yao, R., & Halderman, J. A. (2009). *Analysis of the Green Dam censorware system*. Retrieved from <https://www.eecs.umich.edu/techreports/cse/2009/CSE-TR-551-09.pdf>
- Yang, G. (2009). *The power of the Internet in China: Citizen activism online*. New York, NY: Columbia University Press.
- Yang, G., & Jiang, M. (2015). The networked practice of online political satire in China: Between ritual and resistance. *International Communication Gazette*, 77, 215–231.
- Yuan, E. J. (2012). Language as social practice on the Chinese Internet. In: K. St. Amant & S. Kelsey (eds) *Computer-mediated communication across cultures: International interactions in online environments* (pp. 266–281). Hershey, PA: IGI Global.
- Zhang, C. (2017, July 10). Tuling jiqiren fabu tuling gushiji fangan shendu buju ertong changjing [Turning Robot announces Touring story bot, expanding into children's market]. *IT168*. Retrieved from <http://digital.it168.com/a2017/0710/3143/000003143638.shtml>
- Zhao, C. (2018, May 1). Black mirror in China? 1.4 billion citizens to be monitored through social credit system. *Newsweek*. Retrieved from <http://www.newsweek.com/china-social-credit-system-906865>

Author biography

Yizhou (Joe) Xu is a PhD Student in Media & Cultural Studies at the University of Wisconsin–Madison's Department of Communication Arts. His research interest deals with the software development industry in China, particularly dealing with the roles of state policy, digital labor, and platforms. Prior to the University of Wisconsin–Madison, Joe was a documentarian and broadcast journalist based in Beijing working for new agencies including CBS News, NPR, and Swiss TV. He is the producer behind the feature-length documentary *The People's Republic of Love* (2016).