CARNEGIE
ENDOWMENT FOR
INTERNATIONAL PEACE

FEBRUARY 2024

Tracing the Roots of China's AI Regulations

Matt Sheehan

Tracing the Roots of China's AI Regulations

Matt Sheehan

For your convenience, this document contains hyperlinked source notes indicated by blue-colored text.

Carnegie does not take institutional positions on public policy issues; the views represented herein are those of the author(s) and do not necessarily reflect the views of Carnegie, its staff, or its trustees.

Carnegie Endowment for International Peace
Publications Department
1779 Massachusetts Avenue NW
Washington, DC 20036
P: + 1 202 483 7600
F: + 1 202 483 1840
CarnegieEndowment.org

This publication can be downloaded at no cost at CarnegieEndowment.org.

Contents

Abbreviations

AIIA: Artificial Intelligence Industry Alliance

CAC: Cyberspace Administration of China

CAICT: China Academy of Information and Communications Technology

MIIT: Ministry of Industry and Information Technology

MOST: Ministry of Science and Technology

MPS: Ministry of Public Security

NPC: National People's Congress

NRTA: National Radio and Television Administration

SAMR: State Administration for Market Regulation

Summary

In 2021 and 2022, China became the first country to implement detailed, binding regulations on some of the most common applications of artificial intelligence (AI). These rules formed the foundation of China's emerging AI governance regime, an evolving policy architecture that will affect everything from frontier AI research to the functioning of the world's second-largest economy, from large language models in Africa to autonomous vehicles in Europe.

U.S. political leaders often warn against letting China "write the rules of the road" in AI governance. But if the United States is serious about competing for global leadership in AI governance, then it needs to actually understand what it is competing against. That requires examining the nuts and bolts of both China's AI regulations and the policy process that shaped them. This paper is the second in a series breaking down China's AI regulations and pulling back the curtain on the policymaking process shaping them.

The Chinese Communist Party (CCP) and the Chinese government started that process with the 2021 rules on recommendation algorithms, an omnipresent use of the technology that is often overlooked in international AI governance discourse. Those rules imposed new obligations on companies to intervene in content recommendations, granted new rights to users being recommended content, and offered protections to gig workers subject to algorithmic scheduling. The Chinese party-state quickly followed up with a new regulation on "deep synthesis," the use of AI to generate synthetic media such as deepfakes. Those rules required AI providers to watermark AI-generated content and ensure that content does not violate people's "likeness rights" or harm the "nation's image." Together, these two regulations also created and amended China's algorithm registry, a regulatory tool that would evolve into a cornerstone of the country's AI governance regime.

Contrary to popular conception in the rest of the world, China's AI governance regime has not been created by top-down edicts from CCP leadership. President Xi Jinping and other top CCP leaders will sometimes give high-level guidance on policy priorities, but they have not been the key players when it comes to shaping China's AI regulations. Instead, those regulations have been the product of a dynamic and iterative policymaking process driven by a mix of actors from both inside and outside the Chinese party-state. Those actors include mid-level bureaucrats, academics, technologists, journalists, and policy researchers at platform tech companies. Through a mix of public advocacy, intellectual debate, technical workshopping, and bureaucratic wrangling, these actors laid the foundations for China's present and future AI regulations.

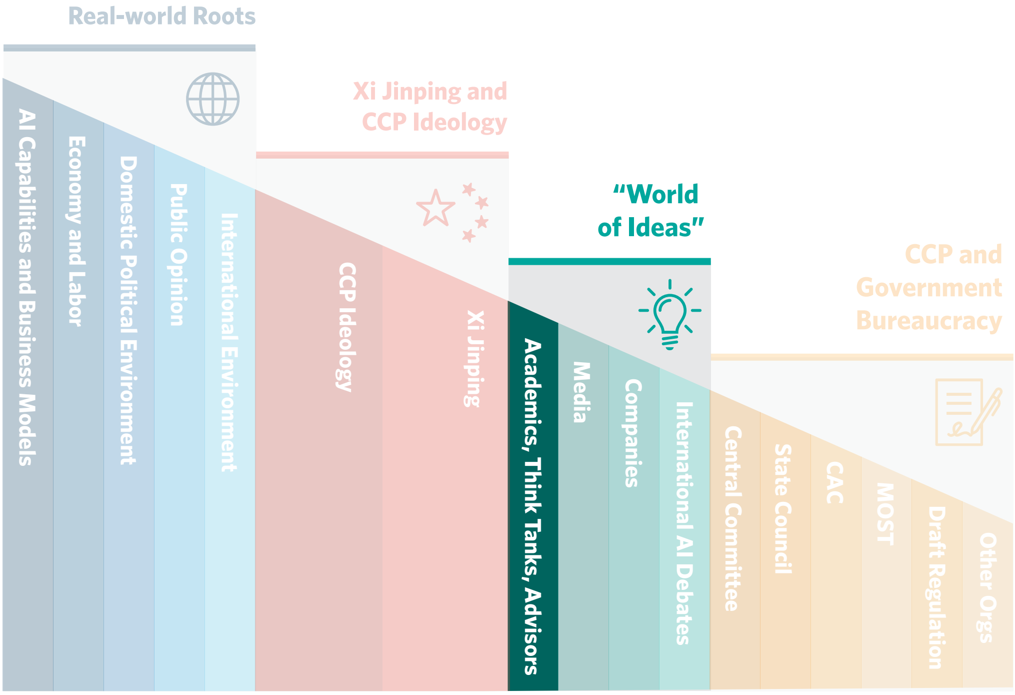
This paper traces the progression of these regulations through the "policy funnel" (see figure 1) of Chinese AI governance. For both recommendation algorithms and deep synthesis rules, the initial spark for the regulation came from long-standing CCP concerns about the creation and dissemination of online content. For the former, the rise of the algorithmically driven news app Toutiao threatened the CCP's ability to set a unified narrative and choose which stories are pushed to readers. In the case of deep synthesis, online face swap videos grabbed the attention of the Chinese public and led government regulators to consider the threat of deepfakes. Over the course of 2017–2020, these concerns made their way through China's bureaucracy. Regulators took a series of stopgap measures in specific applications, while also tasking policy analysts and government-adjacent technical organizations with exploring different regulatory interventions.

At the same time, actors from outside the Chinese party-state were also entering the policy debates, bringing concerns about other social and economic impacts of AI to the table. In 2020, a viral magazine article detailed how food delivery workers were being exploited and put in danger by the algorithms that dictated their routes and delivery times. The piece was built on a foundation of scholarship by Chinese sociologists of labor, and it prompted a massive public outcry against the companies and technologies responsible. Within a year, that outcry translated into provisions on the role of recommendation algorithms in worker scheduling. Chinese technology companies were also actively trying to shape the coming rules, successfully nudging regulators away from using the more ominous-sounding term "deepfakes" and toward the more anodyne version, "deep synthesis."

All throughout this process, Chinese academics, journalists, companies, and even state-run media outlets were actively digesting international AI debates. They would analyze, adopt, and adapt ideas from the United States and elsewhere, covering everything from algorithmic transparency requirements in New York City to deepfake videos created by Buzzfeed News. That willingness to absorb international ideas provides a channel for engagement and even influence by U.S. actors. It also marks a real advantage for Chinese regulators who are willing to learn from and adapt ideas from abroad, regardless of whether they come from a geopolitical friend or foe.

Tracing the genesis of China's AI regulations provides a window into one of the most dynamic and consequential policy areas of today. Understanding domestic Chinese AI governance can clarify the party-state's approach to the technology and to the trade-offs inherent in governing it. Identifying the key actors, institutions, and mechanisms at play also makes it possible to begin constructing a working model of how China makes AI regulations, one that can help predict what might come next.

Figure 1. Exploring the "Policy Funnel" of China's AI Governance



Think tanks, academics & gov advisors

Much public debate on AI governance occurs in publications and workshops featuring academics, researchers, and government advisers. Below are select organizations and individuals shaping these debates.

Think tanks: CAICT; CASTED; Tsinghua I-AIIG

Academics: Zhang Linghan (张凌寒); Liang Zheng (梁正); Xue Lan (薛澜); Zeng Yi (曾毅)

Advisers: National New Generation AI Governance Expert Committee

To use this interactive, please visit CarnegieEndowment.org.

Introduction and Overview

The Foundations of Chinese AI Governance

China is regulating AI, and the rest of the world would be wise to pay attention. Since 2021, the Chinese state has rolled out a series of targeted and binding regulations that constitute some of the first major moves by an AI power to govern one of the most transformative technologies of our time. These regulations target recommendation algorithms, deep synthesis, generative AI, and most recently facial recognition. China is now debating whether to create an overarching national AI law that could be written and rolled out in the years ahead.

This paper is the second in a series of three analyzing China's AI regulations and the forces shaping them. AI is a diffuse technology, and China has begun to introduce policies to address its application in several fields, including autonomous vehicles and medicine. This series of papers analyzes a specific subset of AI-relevant regulations: regulations drafted by the Cyberspace Administration of China (CAC) to deal with public-facing online algorithms and generative AI. The first paper in the series gave an overview of these regulations to date and analyzed their core motivations and structural characteristics. This paper focuses on China's first two significant regulations: the 2021 provisions governing recommendation algorithms and the 2022 regulation on "deep synthesis." The third and final paper in this series will examine the rich debates that shaped China's 2023 regulation on generative AI and what they signal about the road ahead for Chinese AI governance.

Recommendation algorithms are an omnipresent application of AI, powering everything from social media feeds to e-commerce platforms and navigation apps. But they are also an odd target for a country's first AI-focused regulation. Recommendation algorithms do not feature prominently in international AI governance debates because most applications of the technology are seen as relatively anodyne. The regulation on "deep synthesis" provoked similar head-scratching, as the term was unknown outside of China and sparsely used within the country.

Some provisions within the rules themselves also appeared puzzling to outsiders. Alongside some straightforward requirements on recommendation algorithms, such as granting users the right to switch off an algorithm, the recommendation algorithm regulation also contains a grab bag of seemingly unconnected requirements. It bars excessive price discrimination, bans the synthetic generation of fake news, and requires algorithm providers to protect the rights of gig workers. Together the two regulations also established and refined a new algorithm registry, a regulatory tool that has evolved into a cornerstone of China's AI governance framework.

Why did China choose to target recommendation algorithms and deep synthesis? How did these seemingly unrelated requirements get included? And what does that process reveal about the motivations and trajectory of Chinese AI governance? This paper tackles these questions by taking a "reverse engineering" approach (see box 1).

Box 1. Reverse Engineering Chinese AI Governance

This paper is one of a series that attempts to better understand Chinese AI governance by reverse engineering the regulations themselves. The analysis begins with the finished product: the text of the regulations on AI and algorithms that China has already adopted. The regulations are then broken down into their component parts—the terminology, concepts, and requirements embedded in them—and those parts are traced backward through China's four-layered "policy funnel" (see figure 1). By tracing the evolution of China's AI regulations through the policy funnel, this approach aims to both deeply understand China's existing AI regulations and to help predict what new provisions may be coming around the bend.

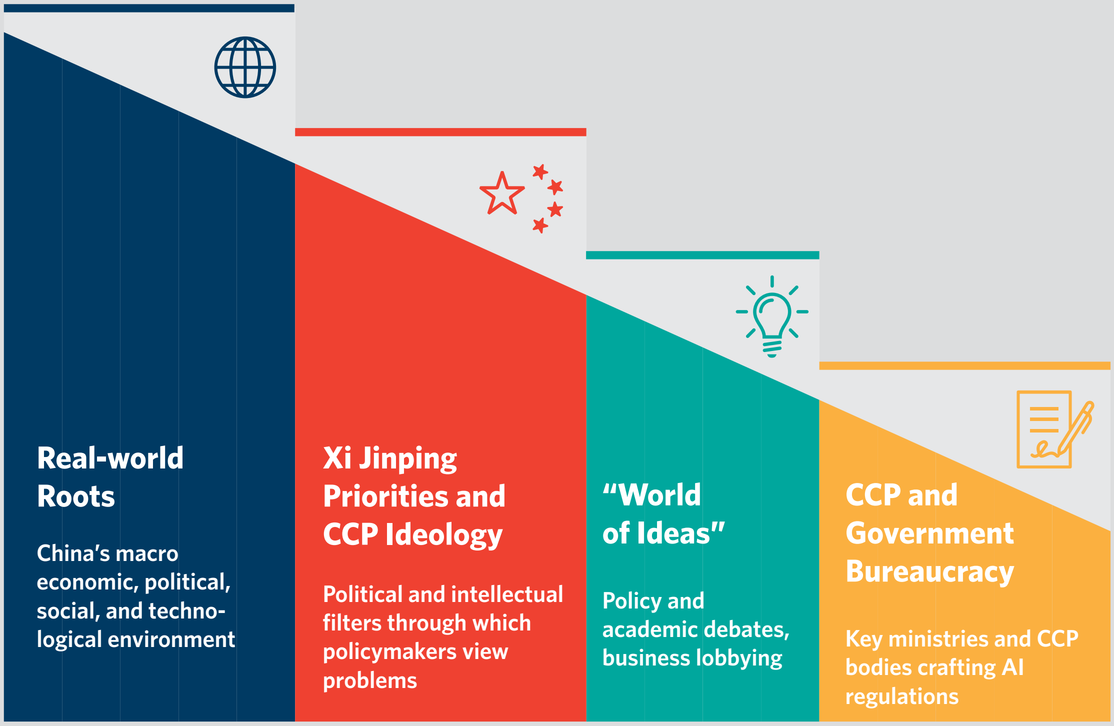
Those four layers of the policy funnel are real-world roots, Chinese President Xi Jinping and CCP ideology, the "world of ideas," and the party-state bureaucracies. Real-world roots include China's macro-level economic, political, social, and technological environment. The forces and real-world events captured in this layer create the need for new policy, but they also limit the options for regulators. The elements of the second layer—Xi's worldview and CCP ideology—act as the political and intellectual filters through which policymakers then understand these issues. To date, Xi does not appear to be actively involved in writing specific AI regulations in China; instead, he sets the direction of travel for policy actors and provides the ultimate backstop for decisions.

The "world of ideas" layer is the most complex and often the most consequential. It is composed of the policy and academic debates that generate new policy proposals, as well as the corporate lobbying that attempts to steer or water down these proposals. While these public debates do not fully dictate policy, they provide the intellectual grist for the bureaucratic mill. The final layer of the policy funnel, the CCP and government bureaucracies, contains the key ministries and CCP bodies that draft and finalize regulations. While the Cyberspace Administration of China (CAC) has been the key actor driving regulation so far, other institutions may take on leading roles as AI governance evolves and expands, including the Ministry of Science and Technology or China's State Council and National People's Congress.

The policy funnel described here and depicted below (see figure 2) is meant to illustrate a conceptual framework for understanding Chinese AI policymaking. In reality, the specific events leading to each new regulation are unique, and the process is often ad hoc. The layers depicted here are porous, and regulations do not proceed through them in a purely linear fashion. Instead, they often pinball forward and backward through these layers, getting shaped and reshaped by different combinations of academics, bureaucrats, public opinion, and CCP ideology. But there is often a pattern to this process, and this paper seeks to provide a grounded and granular look into the underlying forces and key players in that pattern.

For a more detailed examination of these four layers and prominent actors within each layer, see the first paper in this series.

Figure 2. The Four Layers of Chinese AI Governance Policymaking


Real-world Roots
China's macro economic, political, social, and technological environment
Xi Jinping Priorities and CCP Ideology
Political and intellectual filters through which policymakers view problems
"World of Ideas"
Policy and academic debates, business lobbying
CCP and Government Bureaucracy
Key ministries and CCP bodies crafting AI regulations

The following section highlights the key provisions and terminology in these regulations. The paper then traces the history and evolution of these ideas over the past seven years, revealing a dynamic policymaking process. The initial impetus for these regulations came from longstanding concerns of the CCP and the Chinese government over online content. Those concerns about how online content is created and disseminated represent the through line in all of China's early regulations on algorithms and AI.

But as these regulatory proposals gathered momentum, a range of other actors from across Chinese society—academics, journalists, tech companies, and Chinese citizens—reshaped the regulations. Public outcry and intellectual ferment from these groups led to changes in the scope and requirements of the regulations, with new provisions added to address their concerns. Many of these policy actors were absorbing and engaging with ideas about AI governance that were developing in the United States and other countries, constituting an informal channel for both influence and debate.

Tracing the development of these regulations reveals a dynamic policymaking process that weaves together CCP concerns, bureaucratic politics, public discourse, corporate lobbying, academic analysis, and international AI governance debates. Whether one wants to compete against, engage with, or simply understand China's AI governance initiatives, this wider policymaking process must be examined as well.

What's in the Regulation on Recommendation Algorithms?

China's rules on recommendation algorithms are as much about internet regulation as AI. They were initially conceived to target not AI technology in itself but a specific technology application—online recommendation algorithms—that happened to be powered by AI and machine learning. In doing so, they incorporated many components from earlier Chinese regulations on online content, some of which are described below.

The central goal of the rules is to regulate the way in which algorithms affect the promotion and dissemination of content online. Recommendation algorithms presented a new challenge to the Chinese state's long-standing internet controls, and so the CCP singled them out for regulation. In fact, the term "artificial intelligence" (人工智能) does not even appear in the regulation itself.

And yet, the regulation is inextricably linked to AI. The rapid adoption of machine learning algorithms by leading Chinese internet platforms is what created the need for the regulation. And it created precedents and regulatory tools that have become key pieces of China's subsequent AI regulations, laying a foundation for the country's wider governance regime.

Most of the requirements in the algorithmic recommendation regulation can be grouped into three categories: content requirements for service providers, user rights, and business practices (see box 2.). The provisions pertaining to online content reflect the original motivation for the regulation, while those around user rights and business practices appear to have been added on during the policy incubation process.

Box 2. Provisions on the Management of Algorithmic Recommendations in Internet Information Services

Draft: August 27, 2021. Signed: December 31, 2021. **Implemented:** March 1, 2022.

Signatories: Cyberspace Administration of China (CAC), Ministry of Industry and Information Technology, Ministry of Public Security, State Administration for Market Regulation. Draft regulation issued solely by the CAC.

Available in Original Chinese, English Translation.

Key Requirements

Recommendation algorithm service providers must:

- "actively transmit positive energy" and not "disrupt economic and social order"
- create mechanisms for manual intervention in "top searches" and "hot topics" so that promoted content conforms to "mainstream value orientations"

Users have new rights to:

- turn off algorithmic recommendations for an app or website
- select or delete specific user tags for personalizing content recommendations
- obtain an explanation if an algorithm has a major impact on a users' rights

Businesses must:

- not use algorithms for monopolistic or unfair business practices
- not undertake "unreasonable" price discrimination based on user characteristics
- protect the rights to fair compensation and adequate rest for workers whose schedules are set by algorithms

Finally, the regulations creates a **new regulatory tool, the algorithm registry (**算法备案系统). Algorithms that can affect public opinion or "have social mobilization capabilities" must provide basic information on the algorithm and conduct algorithm security self-assessments.

What's in the Deep Synthesis Regulation?

China's second major regulation targets "deep synthesis" (深度合成), a relatively new term that refers to the use of deep learning to synthesize or generate content—what is today called generative AI. The regulation casts a wide net, including the use of algorithms to generate or alter text, images, video, audio, and other virtual renderings, such as the metaverse (see box 3). The term deep synthesis and the wide regulatory scope are an expansion from narrower CCP concerns about deepfakes that sparked the development of the regulation. This paper traces that evolution, revealing both the roots of, and corporate influences on, policymaking.

The deep synthesis regulation aims to guard against the creation of misleading or politically sensitive content, while also protecting the data and privacy rights of Chinese citizens from those using or developing AI. It tries to address these problems at three levels: broad ideological guidance, specific prohibitions and requirements, and the watermarking of synthetically generated content.

Box 3. Provisions on the Administration of Deep Synthesis Internet Information Services

Draft: January 28, 2022. Signed: November 25, 2022. **Implemented:** January 10, 2023.

Signatories: Cyberspace Administration of China (CAC), Ministry of Industry and Information Technology, Ministry of Public Security. Draft regulation issued solely by the CAC.

Available in Original Chinese, English Translation.

Key Requirements

Broad ideological guidance:

- "respect social mores and ethics"
- "adhere to the correct political direction, public opinion orientation and values trends"

Specific prohibitions and requirements for deep synthesis service providers:

- Do not produce, publish, or transmit "fake news"
- Conduct technical or manual reviews of deep synthesis prompts and outputs
- If editing biometric features of a person, remind users to obtain the person's consent
- Conduct in-house or third-party security assessments if either editing biometric features or producing content that "might involve national security" or "the nation's image"

- Requires organizations providing "technical support" for deep synthesis service providers to conduct security assessments and make filings in the algorithm registry

Watermarking requirements for service providers:

- All generated or edited content must contain a digital watermark for identification
- If generated content "might cause confusion or mislead the public," providers must include a "conspicuous label in a reasonable position"
- Note: watermarking requirements were further clarified in a 2023 technical standard

The signing of the deep synthesis regulation on November 25, 2022, marked the end of the first phase of China's regulations on AI and algorithms. The regulations on algorithmic recommendations and deep synthesis had developed in parallel, moving through China's policy funnel and party-state bureaucracy together. These two regulations would lay the foundations of China's emerging AI governance regime, creating new precedents and regulatory tools that would be embedded in subsequent regulations on generative AI and facial recognition. (The term "party-state" refers to the combination of institutions comprising both the Chinese government and the CCP.)

The motivations driving and forces shaping these two regulations are best revealed by looking backward. Tracing the regulations' terminology and provisions back to their origins reveals how CCP concerns, social forces, and intellectual inputs all shaped these regulations.

The Roots of the Regulations

The regulations on recommendation algorithms and deep synthesis emerged separately, but they developed in parallel and were soon bureaucratically woven together, moving through the policy funnel in near unison.

In both cases, the initial spark for the regulation came from party-state concerns about how a new technology application was affecting the creation and dissemination of online content. This core concern provided the impetus to begin the policy process, and it would go on to become the backbone of the final regulation. Understanding how that content-focused backbone came to be offers insight into the CCP's core motivations and process for

AI regulation. The second half of this paper explores how other nongovernmental forces expanded the scope of the regulations to include other social and economic concerns related to the technology.

Why Recommendation Algorithms? The Trouble with Toutiao

The first clue to the origins of the algorithmic recommendation regulation comes in the title itself. "Algorithmic recommendation" (算法推荐) is a sufficiently unique term that it is possible to identify when and in what context it first surfaced in Chinese state media. Those early media discussions of recommendation algorithms reveal how a backlash against algorithm-driven news apps planted the seeds that would grow into one of the world's first major binding regulations on the technology.

The CCP uses China's state- and party-run media outlets as a venue to communicate its priorities and explain its policies. When prominent outlets highlight a problem or promote an industry, that can act as a signal to actors both inside and outside the party. That signal can lead lower-level regulators to turn their attention to that problem, and it can lead businesses to double down on (or withdraw from) that industry. How often a particular term is mentioned in state-run newspapers can serve as a rough proxy for CCP interest in a topic, and the arguments found in those articles are a window into how the party is thinking about that issue. In some cases this media coverage is forward-looking, laying the political and intellectual groundwork for future action. In other cases it is backward-looking, explaining to the public and party members why a certain action was taken. In the case of state media discussion of recommendation algorithms, these early discussions were forward-looking, signaling and sharpening the ideas that would go on to drive policy.

Between 2014 and 2016, the terms "recommendation algorithm" and "algorithmic recommendation" began appearing in some state-run media articles about business and technology. These articles were relatively rare, one every couple of months, and the use of the terms was marginal to the central focus of the article. In late 2016 and early 2017, that began to change. The number of articles referencing algorithmic recommendation increased, and many of them discussed the risks of this new technology. Several of them specifically singled out one popular news app for attack: Jinri Toutiao (今日头条, Today's Headlines).

Toutiao (pronounced roughly: "toe-teeow"), as the app is often called, is an algorithm-driven news aggregator and content platform that delivers a personalized stream of news and other content to users. It was created in 2012 by ByteDance (the company that would go on to create TikTok), and by 2016 it had become China's most popular news app. But just as Toutiao was surging in popularity, ByteDance founder Zhang Yiming stumbled into trouble with Chinese authorities.

In December 2016, Zhang gave a wide-ranging interview to leading business magazine Caijing, discussing everything from his personal life philosophy to how Toutiao dealt with vulgar content. In the conversation, Zhang portrayed Toutiao as a neutral platform whose algorithm's sole goal was to match users with content that they were interested in, regardless of whether that content was high-minded or "vulgar." He explicitly disavowed the idea that his company should try to guide users or inculcate any sort of "values" in them. When asked by the interviewer, he rejected the idea that Toutiao was a media company or the suggestion that it needed an "editor-in-chief."

That laissez-faire approach to content moderation reflected Zhang's own beliefs and showcased how empowered China's leading technology entrepreneurs felt at the time. They were rich, publicly idolized, and often supported by a government that was going all out for "innovation." But the approach also ran counter to a growing push for ideological conformity in Chinese media. Earlier in 2016, Xi had demanded that official media outlets serve the party, saying that the surname of these outlets must be "the party" (媒体必须姓党). In China, nonstate media outlets are restricted from publishing content on a wide variety of public events and government actions.

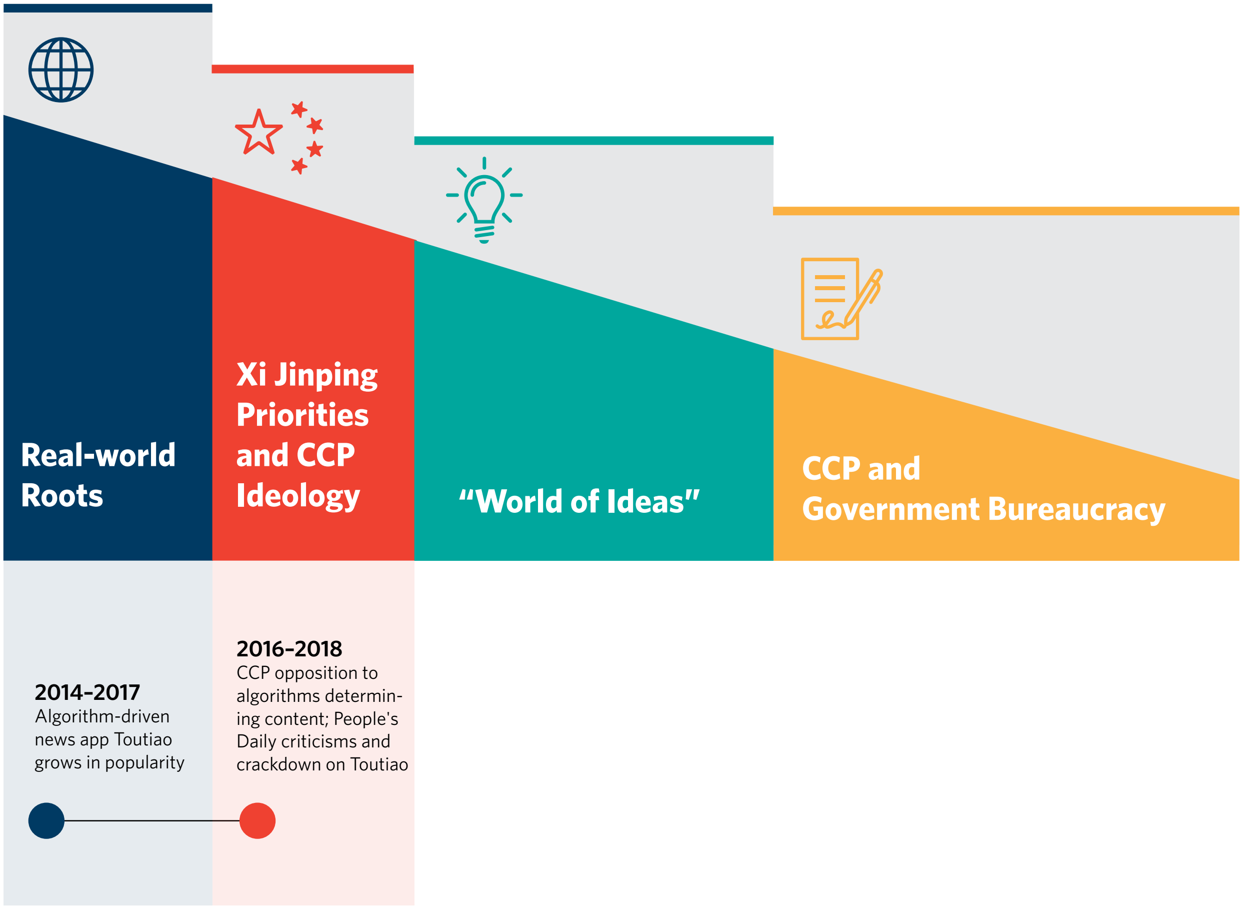
Zhang likely believed his algorithm-driven platform to fall outside of these structures, but the *People's Daily*—the flagship newspaper of the CCP—quickly set him straight. On December 23, 2016, the *People's Daily* ran an editorial that never mentioned Toutiao by name but directly rebutted Zhang's statements. It emphasized the need for proactive "guidance" of public opinion to diffuse social conflicts. It said, "In an era dominated by algorithms, there is an even greater need for 'editors-in-chief' who can guard, guide, and lead."

Over the spring and summer of 2017, criticism of Toutiao in state-run media outlets ramped up, with articles covering everything from the app's alleged copyright infringement to the proliferation of vulgar content. That criticism came to a head in September 2017. From September 18 to 20, the website of the *People's Daily* ran a series of three articles on consecutive days attacking recommendation algorithms, Toutiao in particular. Each of the three articles criticized recommendation algorithms from a slightly different angle.

The first article, titled "Algorithms Cannot Be Allowed to Decide Content," claimed that vulgar content was spreading on these platforms because algorithms had replaced human editors in deciding what content was featured. It once again invoked the need for algorithm-driven platforms to have an "editor-in-chief." The second article criticized recommendation algorithms for trapping readers in an "information cocoon" where they are not exposed to other viewpoints. The third article attacked Toutiao on copyright issues and argued that platforms driven by recommendation algorithms could harm creators and end up producing "the opposite of innovation." It concluded by calling on companies to exercise self-discipline and on the government to improve relevant regulations. With these three articles, the CCP signaled its intense displeasure with the way recommendation algorithms were being deployed in China's media landscape (see figure 3).

In the following months, both official media and regulators began turning up the heat on Toutiao and other algorithm-driven news apps (see figure 3). In December 2017, the Beijing branch of the CAC called managers of Toutiao and another media app to its office to lecture them for spreading vulgar content and operating an online news platform without a permit. Two days later, Toutiao shut down the app's "Society Channel" (a feed for social news) and replaced it with the "New Era Channel" as the default feed for users. The "new era" (新时代) is a political catchphrase popularized under Xi, and the channel promoted news in line with the CCP's narratives. In a public announcement, Toutiao management said the move was made to better "promote the spirit of the [CCP's recent] 19th Party Congress and report on the construction of the new era."

Figure 3. Real-World and CCP Roots of the Recommendation Algorithm


Real-world Roots
Xi Jinping Priorities and CCP Ideology
"World of Ideas"
CCP and Government Bureaucracy

2014–2017
Algorithm-driven news app Toutiao grows in popularity

2016–2018
CCP opposition to algorithms determining content; People's Daily criticisms and crackdown on Toutiao

ByteDance tried to fend off further criticism by holding an event in Beijing in January 2018 on “making algorithms open and transparent.” At the event, Toutiao’s senior algorithm architect explained the fundamentals of how the app worked. But the blows kept coming. In late March, a prominent business show on the state-run broadcaster CCTV ran an exposé accusing Toutiao of violating the law by presenting users with shady medical advertisements. In April, both the CAC and the National Radio and Television Administration (NRTA) separately called in representatives from Toutiao and a short video app, Kuaishou, for more criticism. A few days later, the CAC ordered app stores to block all downloads of Toutiao for three weeks and of a few other apps for shorter periods of time. Not to be outdone, the NRTA then permanently banned another ByteDance app, Neihan Duanzi, that was used for sharing jokes and memes. After the block on downloads, Zhang publicly apologized and pledged to do more to ensure recommendation algorithms would be better aligned with national values.

China’s regulatory apparatus had both identified recommendation algorithms as a target and taken some blunt measures to discipline the companies. ByteDance survived the first regulatory onslaught, but just barely. The struggles of Toutiao following Zhang’s 2016 interview also provided a smaller-scale preview of the punishment that Alibaba-affiliated Ant Group would face—cancellation of its initial public offering—after a defiant speech by founder Jack Ma in late 2020.

What drove the CCP to single out Toutiao and recommendation algorithms? For starters, there were genuine issues with the way the company and other algorithmically driven content platforms were operating. Aside from copyright infringement, deceptive medical advertisements were a real problem. In 2016 a college student died after undergoing shady medical treatments he had found through an ad on Baidu’s search engine, leading to a popular uproar and a government crackdown. There had also been persistent issues with vulgar and borderline pornographic content on Toutiao and other news apps. At the time, an employee of a different algorithm-driven app relayed to this author their team’s strategy for user retention: whenever the app’s monthly users would begin to dip, they would simply inject more sexual content into users’ feeds and the numbers would go back up.

But for all the potential criticisms of Toutiao and recommendation algorithms, one would prove to be core driver of CCP regulation going forward: the role of algorithms in deciding what content gets pushed to Chinese people. Since it took power in 1949, the CCP has largely been able to decide which news stories are featured prominently in the media. This power faced challenges with the rise of nonstate media and then websites, but in both cases the party could always hold a person responsible for the decision to feature a particular story. Toutiao changed that by handing over decisionmaking power to an algorithm.

The power of the app came from the fact that the content it pushed to each user reflected the user’s interests, not the CCP’s. The app further fractured and segmented China’s information environment, undermining the party’s goal of a unified public narrative of events. Zhang’s declaration that Toutiao did not need an editor-in-chief and would not push any particular "values" to users crystallized the conflict. This was unacceptable and required a response, one that would develop over the next few years.

Face Swaps and Deepfakes

In the year following the initial crackdown on recommendation algorithms, Chinese regulators developed a new set of concerns that would eventually evolve into the 2022 regulation on deep synthesis. But to trace that regulation back to its origins, the first step is changing the terminology. "Deep synthesis" (深度合成) is a term that only entered the Chinese AI lexicon in 2020. To follow this regulatory cycle back before 2020, a different term is needed for the specific technology application that first attracted the authorities' interest: deepfakes (深度伪造).

"Deepfake" is a portmanteau of "deep learning" and "fake" that refers to highly realistic images or sounds created using AI. In late 2017, the term entered the English lexicon from the username of a poster on Reddit, who used AI to swap the faces of female celebrities onto the bodies of performers in adult films. The action sparked a public outcry and greater attention in U.S. technology and policy circles to the threats deepfakes posed for individual privacy and mass disinformation.

In China, deepfakes entered mainstream discourse in a similar way: an internet user swapped the faces of female celebrities in videos. In February 2019 an internet user who went by the online moniker "Face Change Bro" (换脸哥) uploaded a short clip from the 1994 television show *Legend of the Condor Heroes* in which he swapped the face of actress Athena Chu out for another actress, Yang Mi. The video immediately went viral on the Chinese internet, sparking debates among users about the implications for copyright protections, actors' incomes, and AI ethics. It also led to a swift backlash. Within two days Face Change Bro had removed the video and apologized, calling the outcry "a warning to all of us." He added, "AI face-swapping should be used in the correct way. Everyone should respect copyright and likeness rights, and devote attention to the technology itself."

In the following months, Chinese public discussion of deepfakes ramped up in both the press and online. Notably, almost every early article in Chinese state media that discussed deepfakes cited research emerging from the United States. The very first article in Chinese state media to mention deepfakes, an August 2018 piece in *Science and Technology Daily*, reported on a project by the U.S. Defense Advanced Research Projects Agency (DARPA) for identifying fake videos. A February 2019 piece in *Beijing Youth Daily* discussing the Face Change Bro scandal cited a video explainer from the Carnegie Endowment for International Peace titled "How Should Countries Tackle Deepfakes?" Other articles cited technical research from New York University, reports by other U.S. think tanks, and even a deepfake video of former president Barack Obama created by Buzzfeed and actor-director Jordan Peele. Despite the two countries' different political systems and growing tensions between China and the United States at the time, Chinese journalists, scholars, and analysts frequently looked to their peers in the United States for analysis of technology trends and regulatory responses.

This early attention to deepfakes was quickly translated into law. In the spring of 2019, a subcommittee of China's legislature was debating amendments to the civil code. In April, two months after the viral face-swap video, it announced amendments to the section on "likeness rights." These amendments explicitly banned the infringement of a person's likeness rights by "using information technology fabrication." The subcommittee said that "some ministries" had raised the issue of deepfakes and asked that the amendments address problems such as "face-swapping." The amendments were formally adopted in June.

The Face Change Bro incident illustrates the Chinese party-state's attention to online discourse and its rapid regulatory response time. But the amendments to the civil code were just a stopgap solution to a more complex problem. Over the following two years, the Chinese state marshaled its own bureaucracy and government-adjacent technical organizations to build a more robust regulatory response to the threat of deepfakes.

Recommendation and Synthesis Merge in the World of Ideas

By early 2019, the CCP had identified both recommendation algorithms and deepfakes as issues in need of attention. While recommendation algorithms posed a threat to China's online content environment, deepfakes first arose as a threat to citizen privacy and "likeness rights." The party-state had taken some initial measures—corporate punishments and tweaks to civil code—but more substantial regulation was needed to achieve its objectives.

Over the course of 2019 and 2020, the Chinese party-state began laying the groundwork for those regulations by galvanizing an array of government-adjacent technical organizations. These organizations form an ecosystem that is a key part of how the Chinese state acquires knowledge and generates touchpoints with industry that help it to understand emerging technologies. Exploring how that government-adjacent technical ecosystem operates gives new windows into how Chinese AI governance is incubated and eventually enforced. Identifying key organizations and committees can also help with forecasting future regulations.

The initial nudge that got these organizations working on recommendation algorithms and deepfakes came from the CAC. Founded in 2014, it quickly rose to a prominent role in Chinese technology policy under its charismatic (and later purged) first leader. The CAC's central mission has always been regulating online content, but through a series of bureaucratic reorganizations and technocratic land grabs it has expanded its remit to include cybersecurity, data policy, and a wide array of issues touched by the internet. The CAC is "one organization with two nameplates," meaning it sits within the CCP bureaucracy but also has an identity within the government. Its exact institutional structure and legal identity are very complex and are the subject of close examination by leading scholars. For the time being, it is enough to know that the CAC is a powerful, content-focused internet regulator, and it has largely driven the first wave of China's AI regulations.

To begin that process, the CAC engaged the Artificial Intelligence Industry Alliance (AIIA). The AIIA was created in 2017 at the behest of several ministries and is de-facto led by the China Academy of Information and Communications Technology (CAICT), an influential technical think tank underneath the Ministry of Industry and Information Technology (MIIT). The AIIA acts as a service provider and coordinator between the alliance's many members: private companies, state-owned enterprises, academic institutions, and the government. In early 2019, the CAC charged the AIIA with creating a new "Cyberspace Technical Committee" (网信技术委员会) that would investigate issues related to AI's role in online content. The new committee was led by the CAICT, in partnership with Zhejiang University and the Institute of Information of the Chinese Academy of Sciences.

Though the operations of the Cyberspace Technical Committee were not highly publicized, a slide deck uploaded to the website of the AIIA in 2020 recapped the group's work over its first year. That slide deck gives a snapshot of how the CAC first began trying to understand and tackle recommendation algorithms and deepfakes.

The committee chose three technical areas to focus on in its first year: recommendation algorithms, deepfakes, and "content security" (内容安全), the latter largely focused on using AI to detect illegal content. For these AI applications, the committee began developing standards, certifications, technical tools for governance, and "industry self-discipline pledges." To do that, it convened technical seminars and workshops that included academia, think tanks, and leading Chinese platform companies.

For recommendation algorithms, the committee focused on industry self-discipline and technical requirements. It drafted an "industry self-discipline pledge" in which companies promised to protect user data, to respect users' right to choose, and to "spread a positive and healthy online culture." The committee also crafted a set of technical requirements for recommendation algorithms in both the news and finance. By 2020, those draft requirements had been turned into an official evaluation and certification offered through the AIIA for "news information domain trustworthy intelligent recommendation systems." The AIIA and CAICT offer evaluation and certification services for many AI applications, ranging from deep learning software frameworks to turnstiles that use facial recognition. Obtaining these types of evaluations and certifications does not offer full legal protection to companies, but it gives them a sense of where regulation is heading, as well as a selling point when marketing their AI-driven products and services to government procurers. Together, these pledges, standards, and certifications were a rough sketch of where the rules for recommendation algorithms were headed.

For deepfakes, the committee focused on gathering information and developing technical tools for detection. In January 2020, it began collecting examples of deepfake technologies. It requested that companies submit descriptions of software for either creating or detecting deepfakes, as well as any relevant recommendations from the companies. In parallel, the committee began drafting its own evaluation methods for deepfake detection technology, and it built a "prototype system for fabricated video management," though details on this are unclear.

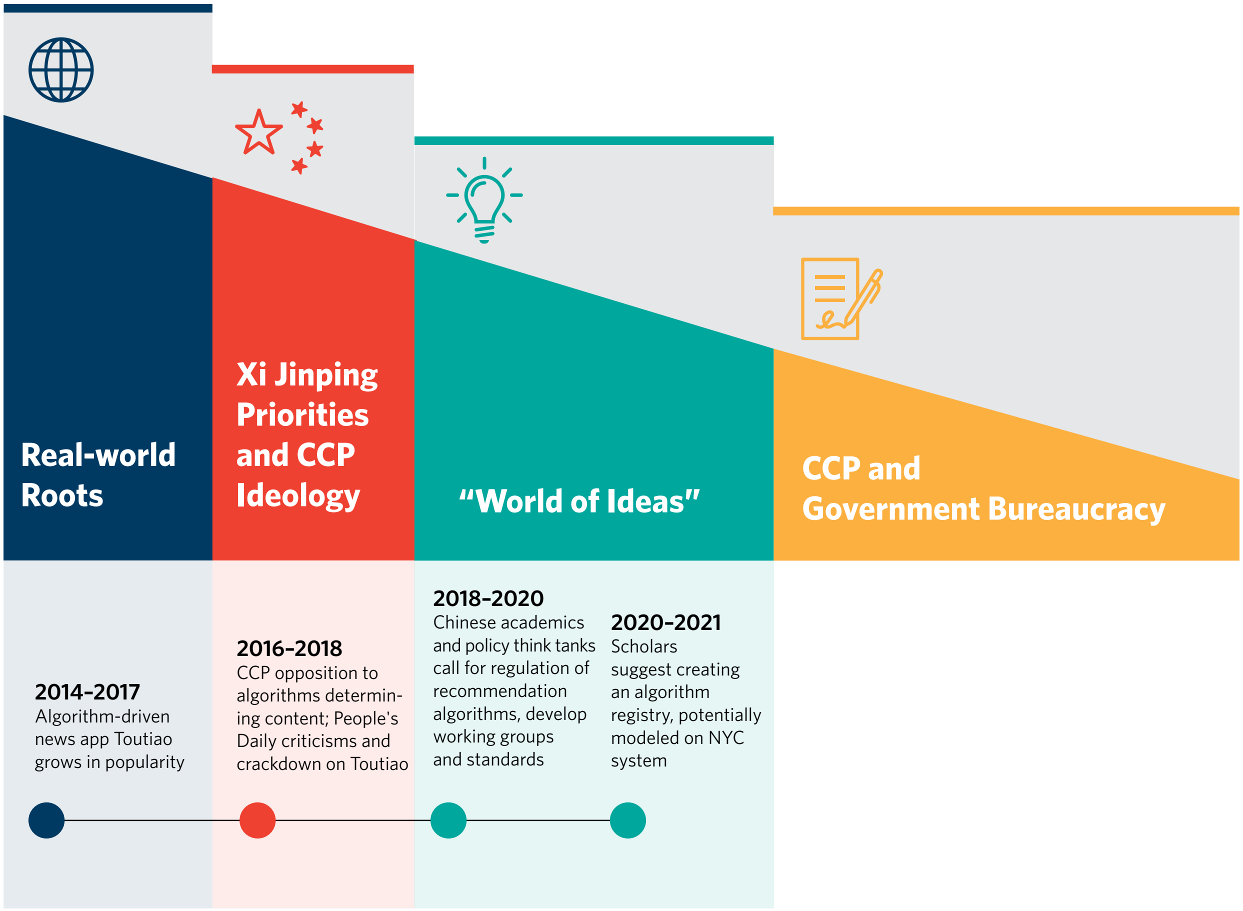
The third area, "content security," focused on AI's role in detecting illegal or harmful content. This work was oriented toward developing AI tools and technical standards to detect undesirable content, not toward developing rules or regulations governing it. As such, this paper does not examine it further.

The committee's work is an example of what the Chinese government calls "collaborative governance" (协同治理). This multistakeholder approach, deployed over the past decade in governing the sharing economy, would have the government avoid simply dictating policy and instead "collaborate with industry associations, platform companies, resource providers, and consumers to address issues that arise" from new technologies and business models. This approach continues today, but the balance between the party-state regulators and other actors is always in flux. From 2015 to 2019, the CAC left much in the hands of companies to self-regulate, but from 2020 to 2022 the administration took a more hands-on—and in some cases, uncompromising—approach to technology regulation.

Exactly what role did the Cyberspace Technical Committee play in the development of China's future regulations on recommendation algorithms and deep synthesis? It offered a chance for the CAC to familiarize itself with the technologies and workshop potential technical and regulatory interventions (see figure 4). For outside observers, the committee's work gave a preview of where regulation was headed and a window into the development of government-adjacent technical certifications. Understanding these committees and technical certifications will offer insight into how the Chinese party-state is likely to approach issues of AI security and safety in the future.

Particularly noteworthy is a new group founded by the AIIA in September 2023, the Safety and Security Governance Committee. (In Chinese, the word 安全 can mean either "safety" or "security" depending on context. In this case, it appears to reference both concepts.) The new committee has been tasked with supporting policy development, building basic infrastructure, and collaboratively exploring issues related to AI safety and security. The new committee is headed by a director at the CAICT, Shi Lin, who also led the earlier Cyberspace Technical Committee covering recommendation algorithms and deepfakes. In December 2023, the committee previewed a new and updated AI risk management framework, which it aims to build on in 2024. Following the work of these types of groups and committees, as well as the development of new standards and certifications for "trustworthy large models," can offer a grounded sense of where regulation in China might be headed.

Figure 4. World of Ideas Roots of the Recommendation Algorithm Regulation


Real-world Roots
Xi Jinping Priorities and CCP Ideology
"World of Ideas"
CCP and Government Bureaucracy
2014–2017
Algorithm-driven news app Toutiao grows in popularity
2016–2018
CCP opposition to algorithms determining content; People's Daily criticisms and crackdown on Toutiao
2018–2020
Chinese academics and policy think tanks call for regulation of recommendation algorithms, develop working groups and standards
2020–2021
Scholars suggest creating an algorithm registry, potentially modeled on NYC system

Writing the Regulations

China's technology regulations tend to be highly iterative. Rules are often rushed out in response to a new issue, and when these prove inadequate the regulators will simply roll out another regulation that clarifies, tweaks, or adds to the requirements in the initial rules. This can appear haphazard, but it is often an intentional strategy that allows regulators to evaluate the effectiveness of different interventions and then double down on the ones that work best. Between 2019 and 2023 this pattern played out in China's AI regulations.

Early Attempts at Rulemaking

In late 2019, when it was overseeing the AIIA's work on deepfakes and recommendation algorithms, the CAC also released two regulations that addressed aspects of these problems. In November 2019, the CAC released the Provisions on the Management of Online Audiovisual Services, which targeted deepfakes. China's Ministry of Culture and Tourism and the NRTA co-signed the regulation. A month later came the Provisions on the Governance of the Online Information Content Ecosystem, a sweeping, content-focused regulation that also addressed online recommendation systems. These two sets of rules were the CAC's first stab at formally regulating these technologies.

The audiovisual regulation banned the use of deep learning to create or spread "fake news," and it required conspicuous labels to be applied when deep learning is used to create "false audiovisual information." The content ecosystem regulation was much broader and largely focused on sorting content into broad categories: "positive" (encouraged), "negative" (discouraged), and "illegal" (banned). It then applied these categories to the use of personalization algorithms, requiring that providers use those algorithms to "actively present" positive information, to "prevent and resist" negative content, and to not display illegal content. The regulation also required that providers "optimize" their recommendation systems and use manual interventions so that lists of "hot topics" and "top searches" display politically desirable content. These two 2019 regulations gave a relatively crude preview of many requirements that would eventually make their way into the more robust regulations of 2021 and 2022.

Why did these regulations come out when they did, and why were they ultimately deemed insufficient?

In the case of the audiovisual regulations, the provisions were likely a slapdash response to the public outcry over a popular new face-swapping app, Zao. The app debuted in August 2019, and within days it was the most-downloaded app in Apple's Chinese app store. In the weeks that followed, online discussion and news articles criticized Zao and discussed the risks to user privacy and misinformation resulting from deepfakes. The CAC's deputy director fielded questions from journalists about regulating the app. The audiovisual regulation was released two months later. One analysis reposted on its website praised the CAC for "conforming to the urgent demands of the broad masses of people" in its adherence to the principle of "urgent needs come first" in regulation. But what the regulation gained in speed, it lost in specificity. The scope of applications covered was vague, perhaps intentionally so, as were many of its requirements, such as those for labeling AI-generated content.

In the content ecosystem regulations, the rules addressing recommendation algorithms were not the central focus. The provisions targeted one related concern—what content gets promoted on homepages and on lists of "hot topics"—but they largely focused on enumerating types of undesirable content. The regulation requires platforms to respect users "right to make selections," but it didn't clarify what that right meant in practice. Getting clarity would require the creation of a regulation that was wholly focused on recommendation algorithms.

One month after the content ecosystem regulation was signed, COVID-19 swept through the city of Wuhan and then China and the world. The massive upheaval caused by the virus dramatically slowed down regulatory action on these topics during the first half of 2020. But by the end of the year, the highest levels of the party turned their attention back to these topics.

In December 2020, the powerful Central Committee of the CCP issued a document intended to guide much of the legislation and regulation over the next five years: the Implementation Outline for the Establishment of a Rule of Law-Based Society (2020–2025). The document contained some overarching guidance and a massive laundry list of specific issues. Buried within a section on internet law issues that needed to be addressed were two AI applications: recommendation algorithms and deepfakes.

The inclusion of these two marked a success for the CAC's policy entrepreneurship. Between 2017 and 2020 the CAC had successfully established itself as the go-to regulator for both topics, and it had managed to get them elevated to the level of the Central Committee. It also meant that more work was needed to hammer down these topics. Over the course of 2021 and 2022, the CAC would lean on Chinese academics, think tanks, and companies to get that work done.

How to Draft a Chinese AI Regulation

Between August 2021 and January 2022, the CAC went on a spree of releasing both drafts and finalized regulations on algorithms and AI. In August and December of 2021, the CAC released a draft and then final version of its regulation on "algorithmic recommendation." The final version was co-signed by the MIIT, Ministry of Public Security (MPS), and State Administration for Market Regulation (SAMR). Between these dates, the CAC and eight other CCP and government departments also jointly released guidance on "strengthening overall governance" of online algorithms. (This document provides high-level principles for regulation of algorithms and will not be covered in this paper.) In January 2022, the CAC released a draft of the new regulation on "deep synthesis," which it finalized ten months later in November, in conjunction with the MIIT and MPS.

How did the policy ideas in these regulations move from the amorphous intellectual ferment of Chinese policy discourse—the "world of ideas" layer of the policy funnel—into concrete requirements in regulations?

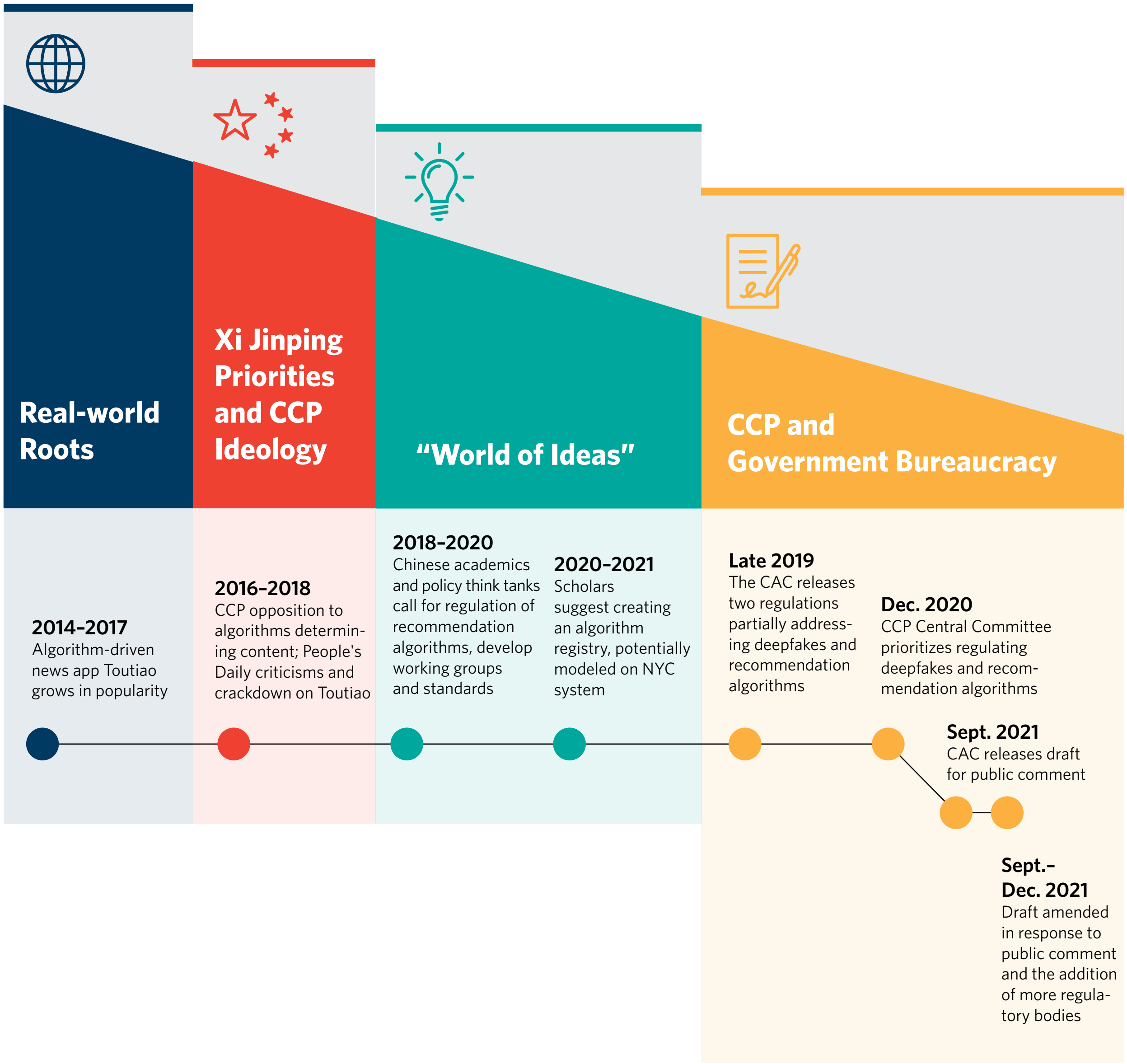
Most of this process happened behind closed doors, with limited windows for outsiders to peek inside. But some characteristics of the processes at play are apparent. The CAC has both power and experience with controlling online content but more limited in-house expertise when it comes to the technical and legal specifics of regulating AI. So it relies heavily on extensive external consultations with legal and technical experts from across academia, think tanks, and industry.

These consultations occur over various stages of the drafting process and range from informal exchanges to formal workshops and specially commissioned reports. A single regulation may see over one hundred different experts consulted. Some of those outside advisers will simply offer general ideas or feedback, while others may be asked to take the pen and draft pieces of the regulation. In some cases, the regulating organization will create a formal "expert drafting small group" ("专家起草小组" or "专班"), while in other cases the process is conducted more informally. The membership of these advisory groups is not always publicized, but some scholars will publicly tout their role as advisers. The CAC also occasionally announces the appointment of groups of legal advisers to the organization as a whole.

Which scholars and institutions are consulted will depend partly on which bureau leads the drafting of a regulation. The CAC is home to over a dozen bureaus, each with a different specialization. The head of a given bureau may have a preexisting network of experts they rely on or a tendency to lean more heavily on technical or legal experts. While the final text of regulations will not list which bureau led the drafting process, the leadership can sometimes be inferred based on which bureau is designated to receive feedback on the draft regulation. Both the recommendation algorithm and deep synthesis regulations appear to have been drafted by the Bureau of Cyber Rule of Law (网络法治局) of the CAC. The draft generative AI regulation appears to have been drafted by the Cybersecurity Coordination Bureau (网络安全协调局).

Once a regulation emerges from a bureau at the CAC, it goes through several layers of internal bureaucratic consultation (see figure 5). This process is not identical for all regulations, but it appears to have been followed for the two regulations covered in this paper. First, a regulation is circulated to the other bureaus within the CAC for feedback, then to other relevant ministries and agencies across the government. Once it has gone through these two rounds of internal consultation, a draft is published to solicit public feedback, including from outside experts, companies, and regular citizens. A revised version is then circulated for feedback once more among relevant ministries and agencies before a final draft is published. While the first drafts are often issued solely by the CAC, the final versions are sometimes co-issued by other ministries or agencies who oversee aspects of the relevant topics (see figure 5).

Figure 5. CCP and Government Roots of the Recommendation Algorithm Regulation


Real-world Roots
Xi Jinping Priorities and CCP Ideology
"World of Ideas"
CCP and Government Bureaucracy
2014–2017 Algorithm-driven news app Toutiao grows in popularity
2016–2018 CCP opposition to algorithms determining content; People's Daily criticisms and crackdown on Toutiao
2018–2020 Chinese academics and policy think tanks call for regulation of recommendation algorithms, develop working groups and standards
2020–2021 Scholars suggest creating an algorithm registry, potentially modeled on NYC system
Late 2019 The CAC releases two regulations partially addressing deepfakes and recommendation algorithms
Dec. 2020 CCP Central Committee prioritizes regulating deepfakes and recommendation algorithms
Sept. 2021 CAC releases draft for public comment
Sept.–Dec. 2021 Draft amended in response to public comment and the addition of more regulatory bodies

New Additions and External Influences

Previous sections of this paper have laid out the roots of these regulations in CCP and state concerns over online content and shown how those concerns made their way through party and state apparatuses. This section looks beyond those content-centric concerns and focuses on key additions to these regulations during the process of drafting and revising them. Tracing the origins of these changes shows how the regulations were shaped by nongovernmental actors from throughout China's broader social, economic, and intellectual ecosystem.

What's New in the Regulations? Going Beyond Content Concerns

The regulation on recommendation algorithms saw some of the most significant additions between the draft and the final version. A provision barring the use of algorithms for unfair competition or monopolistic practices was added. This is likely directly attributable to the addition of the SAMR as a co-signatory of the final version. The SAMR, China's top antitrust authority, took a series of anti-monopoly and unfair competition actions against China's leading platform companies in 2020–2022, culminating in changes to the Anti-Monopoly Law and Anti-Unfair Competition Law. This campaign followed Xi's pledge to build "common prosperity," a broad push to narrow the wealth gap that helped fuel a range of crackdowns, several led by the SAMR, on large corporations. The provision on algorithms' role in reinforcing monopolies was most likely added at its behest, a manifestation of the kind of bureaucratic bargaining that takes place as the CAC seeks buy-in and co-signatories among other agencies.

More mysterious was the addition of Article 20 of the regulation. That article requires recommendation algorithm providers that offer "work dispatch services" (工作调度服务) to protect the rights and interests of workers, including their legal rights to compensation and rest. It also demands that providers improve any algorithms used to assign orders or to dictate workers' salaries and scheduling. This provision stands apart from the rest of the regulation in its focus on worker rights. It has no clear connection to the core concerns of the CAC or its fellow regulators. So where did this issue come from, and how did it make its way into the regulation? The following section will trace its origins.

The most notable changes in the deep synthesis regulation are its title and scope. The roots of the regulation are found in the party-state's concern over deepfakes (深度伪造). Prior documents, including the December 2020 directive from the CCP Central Committee, explicitly used the term "deepfakes" in its call for regulation on the topic. But somewhere between that December 2020 directive and the January 2022 release of the draft regulation, the CAC changed both the terminology and the scope of the regulation to "deep synthesis" (深度合成). In the regulation, deep synthesis is defined as the use of "generative synthesis algorithms" such as deep learning to generate or edit text, images, audio, video, or other "virtual scenes." The term "deepfake" is not in the regulation itself, only appearing in an accompanying article explaining that the regulation responds to the Central Committee's call for addressing deepfakes. Why did the CAC change both the key term and the overall scope of the regulation? This will be answered in a subsequent section.

A final, notable addition to both the recommendation algorithm and deep synthesis regulations is the creation of the algorithm registry (算法备案系统). Both regulations require that providers whose algorithms have "public opinion properties or capacity for social mobilization" submit a filing on their algorithm within ten business days of operation. What that filing should contain is left vague in the text of the original regulation, but the algorithm registry has gone on to become a key component of China's AI governance regime. The origins and mechanics of the algorithm registry are explored further below.

To understand where these additional regulatory requirements came from, this analysis will turn back the clock to 2020 and widen the field of vision beyond the machinations of China's official bureaucracies.

How Sociologists and Investigative Journalists Shaped a Regulation

There was a clear inciting incident behind the algorithmic recommendation regulation's protections for workers: the public outcry in 2020 over the plight of food-delivery workers whose schedules and routes are set by algorithms. That outcry followed the publication of a magazine exposé on the struggles of these workers that went viral, an article that was in turn built on the work of Chinese and American sociologists and anthropologists. Tracing the progression of these ideas from academic papers to a magazine article to public outcry to the party-state's response showcases alternative routes to policy influence in China.

Over the past decade, the restaurant industry in Chinese cities has been transformed by food delivery apps. By 2020, China's food-delivery employed 6 million delivery drivers who were zooming across cities on electric bikes and scooters. Two companies—Meituan and Ele.me—dominate that industry, combining to control 98 percent of the market.

Both companies rely on machine learning algorithms to assign deliveries, set expected wait times, decide delivery routes, and determine driver compensation for each trip. Amid the fierce competition between them, those algorithms were often tuned to reduce both delivery times and driver compensation to an absolute minimum. Workers found themselves scrambling to meet these new demands, while struggling to make the deliveries needed to earn a livable income. Many of these workers are migrants from the Chinese countryside who lack both proper contracts and any social safety net.

Outrage about the treatment of these workers erupted in September 2020 following the publication of a long-form magazine exposé titled "Delivery Workers, Trapped in the System" (original Chinese). The piece was published by *Renwu* ("People") magazine, an outlet under a state-owned publishing house. But *Renwu* is known for some edgy publication decisions, including an interview with an early coronavirus pandemic whistleblower that was later censored. The article about delivery workers detailed how the unreasonable delivery times and low compensation set by Ele.me and Meituan's algorithms forced delivery drivers to run red lights and ride against traffic in their desperation to make a living. It described in gory detail the way these pressures led to the death of drivers:

> One afternoon this spring, Wei Lai and another rider wearing the same color clothing as him were stopped at an intersection waiting for the light to change. There were only a few seconds before the light turned green, but the other rider darted into the intersection. At the same time, a fast-moving bus sped through and the rider and his scooter went flying. He died on the spot. Wei Lai said he saw the badly mangled body in the middle of the road, but he did not stop at all. His own order was late. At that time, another order came in and the familiar female voice of the app's delivery assistant chimed in: "Order! From 'Point A' to 'Point B,' please respond after the beep to accept."

The article also explored the technical and sociological underpinnings of the industry, drawing extensively from the work of Chinese sociologists of labor, including Sun Ping from the Chinese Academy of Sciences and Zheng Guanghai from Central China Normal University. These researchers had spent years surveying delivery workers and developing sociological concepts to describe the interplay of algorithms, platforms, workers, and consumers. Those academics in turn frequently cited and built on the work of American academics like Nick Seaver, who developed the concept of "algorithms as culture." Weaving all of these strands together, the *Renwu* article was a remarkably in-depth yet accessible investigation into the role algorithms played in the exploitation of gig workers.

Despite its length and heavy subject matter, the article quickly went viral, dominating social media feeds and sparking intense public backlash against Ele.me and Meituan. The companies attempted to placate these concerns with minor changes to their apps. Ele.me gave users the option to "wait five more minutes" for their delivery, while Meituan added eight minutes of "flexible delivery time." But many commentators dismissed these as cosmetic changes that just shifted the burden onto consumers. Chinese state media piled on the criticism. One leading television anchor demanded the platforms treat their employees as "people, not machines" and called for increased government oversight of the industry. Newspapers associated with different parts of the party-state bureaucracy proposed different solutions, ranging from technical interventions in the algorithms to resetting the contracting relationship between firms and workers.

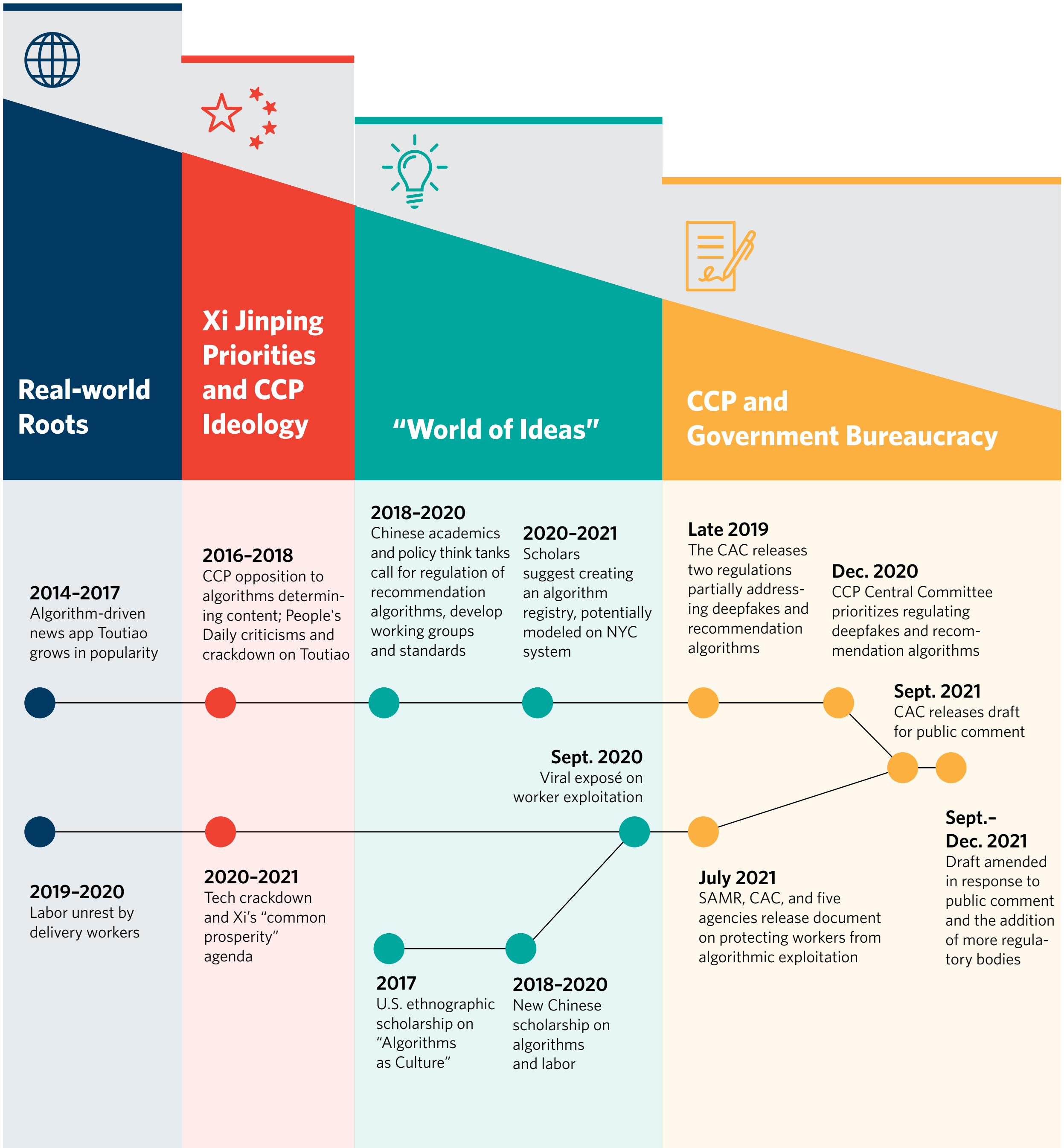
The government's regulatory response came the following summer. On July 7, 2021, the State Council issued a statement on protecting the rights of workers in the platform economy, which included a call for delivery platforms to improve the algorithms used to assign orders and to determine worker compensation. Later that month, the SAMR and the CAC led a group of seven agencies in co-issuing new guiding opinions to "safeguard the rights and interests of food delivery workers." The first provision in the document required platforms not to use "the strictest algorithm" but instead to adopt a "moderate algorithm" in determining the distribution and time expectations for orders. Other provisions in the regulation target issues highlighted in the *Renwu* article, such as setting floors under worker income and increasing enrollment in insurance programs.

In August 2021, the CAC followed up with its draft regulation on algorithmic recommendations, which included the provisions calling for the protection of workers whose schedules and compensation are determined by algorithms. During author discussions with advisers involved in drafting the regulation, those advisers explicitly cited the *Renwu* article and the outcry as the spark that led to the inclusion of these provisions. Among regulatory bodies, the CAC is particularly attuned and responsive to public outcry, a sensitivity that stems from its role monitoring and censoring online discussion (see figure 6).

Though the requirements in both of the regulations are relatively vague—"protect workers' rights and interests"—the companies responded. In July 2022, an engineering manager at one of the delivery platforms told the author that augmenting the delivery algorithm became a top engineering priority during that time. Some of that work can be seen in the public-facing versions of the companies' algorithm registry filings. Both Meituan and Ele.me took pain to highlight how they had loosened up time requirements, with Meituan claiming that its algorithm generates four estimates for delivery time and it chooses the longest one. Whether those changes materially improved the experience of delivery drivers is beyond the scope of this paper, but anecdotal evidence suggests that these workers still face grueling and precarious work conditions.

Despite the Chinese state's eagerness to protect workers from the excesses of algorithms, it was not receptive to direct agitation by those workers. Strikes by delivery workers had been ramping up in the years prior to the *Renwu* article, and the leaders of these strikes were often detained by police. Shortly after one Ele.me delivery driver self-immolated over unpaid wages in January 2021, a prominent labor organizer was arrested. This combination of police crackdowns and new regulations is part of the party-state's playbook for responding to social issues: arrest those who organize unsanctioned collective action and simultaneously enact policies intended to address the underlying cause of public anger.

Figure 6. Roots of the Recommendation Algorithm Regulation



Real-world Roots
Xi Jinping Priorities and CCP Ideology
"World of Ideas"
CCP and Government Bureaucracy

2014–2017 Algorithm-driven news app Toutiao grows in popularity

2016–2018 CCP opposition to algorithms determining content; People's Daily criticisms and crackdown on Toutiao

2018–2020 Chinese academics and policy think tanks call for regulation of recommendation algorithms, develop working groups and standards

2020–2021 Scholars suggest creating an algorithm registry, potentially modeled on NYC system

Late 2019 The CAC releases two regulations partially addressing deepfakes and recommendation algorithms

Dec. 2020 CCP Central Committee prioritizes regulating deepfakes and recommendation algorithms

Sept. 2021 CAC releases draft for public comment

Sept. 2020 Viral exposé on worker exploitation

2019–2020 Labor unrest by delivery workers

2020–2021 Tech crackdown and Xi's "common prosperity" agenda

July 2021 SAMR, CAC, and five agencies release document on protecting workers from algorithmic exploitation

Sept.–Dec. 2021 Draft amended in response to public comment and the addition of more regulatory bodies

2017 U.S. ethnographic scholarship on "Algorithms as Culture"

2018–2020 New Chinese scholarship on algorithms and labor

Note: The policy funnel depicted above is intended as a conceptual framework for illustrating the development of China's regulation on recommendation algorithms. The events are presented in roughly chronological order from left to right, but many of them overlapped and the connecting lines do not necessarily represent a causal relationship. For an interactive breakdown of the sub-components of each layer of the policy funnel, see Figure 1.

From Deepfakes to Deep Synthesis

While worker protections stemmed from investigative reporting and public outcry, the use of the term "deep synthesis" had a different origin: corporate thought leadership.

By early 2020, the party-state had clearly singled out deepfakes as a target for regulation. For a major Chinese technology corporation with big ambitions for AI-generated content, there was reason to worry that regulatory overreach would preemptively smother a whole new AI-driven industry.

The Chinese tech giant Tencent is best known as the creator of WeChat, the Chinese super-app that allows users to chat, make payments, book a doctor's appointment, and hundreds of other functions. But Tencent is also a digital media empire consisting of streaming platforms, a movie studio, and some of the most successful online games ever created. The opportunities to leverage AI across this portfolio are massive. Looking at both the commercial opportunities and the potential regulatory headwinds, Tencent's in-house policy think tank conducted some AI policy thought leadership.

In May 2020, Tencent Research Institute (腾讯研究院) launched a report titled, "AI-Generated Content Development Report 2020: Deep Synthesis's First Year of Commercialization." The report deftly reframed discourse about AI-generated media, switching both the terminology and the focus of the discussions. It argued that using the term "deepfake" to describe all AI-generated content was "amateur and unscientific" and that it unduly stigmatized the technology. Instead it proposed "deep synthesis" as a more neutral and inclusive term. The report then laid out the potential applications of deep synthesis across industries ranging from entertainment to education to healthcare. It predicted that 2020 could be the year that commercial applications of the technology flourished.

The report did not deny that deep synthesis presented new problems nor dispute that some regulation may prove necessary. It described the sordid origin of the term "deepfake" on Reddit and described threats posed from fraud, misinformation, and even election interference. A substantial section of the report outlined U.S. and European legislative proposals targeting deepfakes, as well as China's own regulations. It discussed technical countermeasures like detection techniques and industry attempts at self-regulation.

But the report warned against regulatory overreaction. It claimed that deep synthesis "won't erode truth in society, and it isn't a threat to world order." It cautioned that governments should not rush to regulate deep synthesis but instead adopt a tolerant and prudent approach that does not obstruct useful applications and technological innovation.

These conclusions were convenient for the interests of Tencent, but the arguments in the report were not without basis. At the time, much of China's policy discourse was using the loaded term "deepfake" to refer to a wider batch of AI applications, and "deep synthesis" was a more neutral and appropriate term for AI-generated content. (The term "generative AI" had not yet entered mainstream policy discourse.) As with much of the work from Tencent Research Institute, the report was well-researched, technically sophisticated, and forward-looking. It was also an attempt to steer the policy conversation away from the threat of deepfakes and toward things the Chinese party-state wants, like innovation and increased productivity.

On one level, this appears to have worked. Prior to Tencent's report, the term "deep synthesis" had never appeared in Chinese state media. In the eighteen months following the report's publication, the term gained some traction in policy discourse, and in January 2022 it jumped to the center of that discourse when the CAC published its draft regulation. How exactly the term entered and held sway within the drafting process remains a mystery. It is possible that Tencent's government relations team directly pushed for its adoption or that it simply entered the Chinese policy lexicon and then was taken up by regulators.

But on another level, the report may have partially backfired for Tencent. Chinese regulators had previously been focused explicitly on deepfakes, most often in the form of face swapping. In an attempt to showcase the many productive uses of the technology, Tencent's report widened the aperture to include many other kinds of AI-generated media that would not properly be called deepfakes. Regulators ended up using not only the term but also the much more expansive definition in the regulation, including things like AI-generated text and image enhancement. It is possible that regulators were already planning to cast a wide net on the promised deepfake regulation and that the report did not actively expand that scope. But it is also possible that this bit of corporate thought leadership contributed to a more far-reaching—if somewhat less severe—regulation.

Roots of the Algorithm Registry

The most consequential feature of the two regulations was the creation of the algorithm registry (算法备案系统, sometimes translated as "algorithm filing system"). This system went largely unremarked upon at the time, but it has gone on to become a key piece of China's AI governance toolkit—a way for the party-state both to gather information and to impose requirements on Chinese AI developers. So, what is the registry, and where did it come from?

China has a long tradition of requiring companies and individuals to register projects, products, and much more with the government. These registration systems (备案系统) keep track of everything from construction projects to published books to the color of each car. In most cases, they only require basic information: the individual's name, the company or organization, and some information specific to the object or activity. They are intended simply to create a record for the government and are explicitly not licensing systems—though, in the case of the algorithm registry, this line is becoming blurred.

After the dawn of the internet, the Chinese government began creating registries for cyberspace, beginning with noncommercial websites (as commercial sites require a license). These early online registries were created and managed by the bureau that would go on to become the MIIT. But the CAC entered the fray in 2019 by creating a registration system for blockchains. Given this background, it is not entirely surprising that the CAC would eventually create a registry for algorithms, but it also was not a given. These registries are usually created only after a new regulation demands them. For example, China did not have a formal registry for mobile apps until August 2023, when a new law covering online fraud mandated it.

China's recommendation algorithm regulation created that mandate. It required that all providers whose algorithms have "public opinion properties or the capacity for social mobilization" submit a registration within ten days of operating. The information explicitly mandated by the regulation is fairly basic: name and type of algorithm, form of service, application setting, and an "algorithm self-assessment report." When the first batch of registrations were made public in August 2022, they included some additional details, such as a very high-level and nontechnical description of the functioning of the algorithm. These details, and the fact that they were released publicly, were new additions compared to previous registries. If this were the full extent of information given to Chinese regulators, it would provide them with few meaningful insights into the algorithms.

But the CAC was publicizing only a fraction of the information it gathered through the registrations, with much more being kept private due to concerns over trade secrets. Some insight into that additional information is available in a downloadable user manual that the CAC posted on the algorithm registry website. The document is intended to guide algorithm providers through the registration process, and it contains screenshots of various stages of the process. These screenshots show that the CAC is asking for significantly more information on input data, such as whether it contains biometric features or personal identity information. Providers are also asked to list all the open-source and "self-built" datasets that the algorithm was trained on, as well as to upload an "algorithm security self-assessment report." The user manual does not contain any further details on the contents of that report, but these reports are likely modeled on a 2018 document from the CAC on conducting security assessments for internet services with "public opinion properties or social mobilization capabilities."

Many Chinese companies were initially confused as to the scope of applications covered and how to satisfy certain requirements. The first batch of registrations included everything from search-engine filtering algorithms to apps that recommend parenting tips and child-rearing content. And there remained a huge knowledge gap between companies and the CAC regulators when it came to the functioning of the algorithms. The *Wall Street Journal* reported on a meeting between representatives from ByteDance and the CAC, in which officials from the agency "displayed little understanding of the technical details" of algorithms, forcing company representatives to "rely on a mix of metaphors and simplified language."

But both the algorithm registry and the regulators involved have evolved. When the deep synthesis regulation was finalized in December 2022, the registry was updated to include a separate registration process for technical support teams, which face new requirements under the regulation. The biggest changes came with the explosion of generative AI applications in 2023. The new generative AI regulation from July 2023 did not explicitly change the registration process and simply mandated that providers fulfill their existing obligations under the algorithmic recommendation and deep synthesis regulations. But in practice regulators began treating the registration process more like a licensing regime than a simple registration process. They did this by withholding their official acceptance of registrations until they felt satisfied with the safety and security of the models.

This process of ensuring that models are "safe enough" for release began as an informal back-and-forth between companies and regulators. But in October 2023 a key Chinese standards body released a draft standard that laid out specific tests a generative AI model must pass before registering. If implemented, that draft standard would mandate that each dataset used for training contain less than 4 percent objectionable content. For sensitive prompts—covering ideology, politics, race, and gender, among other topics—over 90 percent of the model's outputs must be acceptable. The draft standard includes detailed requirements for several other tasks, such as what percentage of user prompts on sensitive topics a model can refuse to respond to.

Over the past two years, Chinese regulators have repeatedly iterated on and expanded the function of the algorithm registry. It has gone from being a straightforward tool for limited algorithmic transparency to a flexible, multipurpose tool of AI governance.

So where did the idea for creating a new registry for algorithms come from? The earliest public call for the creation of an algorithm registry appears be in 2016. At the time, the Chinese search engine Baidu was embroiled in controversy. As mentioned above, a Chinese student had died after undergoing bogus medical treatments, which he had found through an ad on Baidu, leading to a major public outcry. At a forum with internet policy experts debating the issue, legal academic Liu Deliang advocated for the creation of an algorithm registry, one that would include Baidu's search engine algorithm. Liu argued that without a registry, the public and regulators had no way of knowing whether Baidu had tampered with the search results for financial gain. In the following years, Liu continued to argue for greater regulation on algorithms, including in a 2019 interview with the influential newspaper *Legal Daily* calling for mandatory public disclosure of corporate algorithms to facilitate societal oversight.

But it is also possible that inspiration for China's algorithm registry came from a far different source: New York City. Chinese academics—and to a certain extent, regulators—frequently engage with their counterparts abroad, discussing how different countries have dealt with issues China is now facing. These comparative policy studies go back many decades and span dozens of countries, with economists from Eastern Europe and industrial planners from Japan deeply engaging with and influencing leading Chinese scholars as they confront these policy areas.

During the CAC's consultations with policy experts for the algorithmic recommendation regulation, one expert pointed to New York City's newly created directory of algorithms used by city agencies. The directory was the product of an executive order by the mayor requiring that agencies disclose their use of algorithms in decisionmaking. It includes basic information on the purpose and functioning of the algorithms, including the types of data they were trained on. For example, it describes a "growth model" used by the New York State Education Department to assess teacher performance. The model compares student achievement to projections made by the model, which are based on economic disadvantage indicators, disability, prior achievement, and other factors.

The New York City directory bears many similarities to the public-facing versions of China's algorithm registrations. Both include a basic description of the purpose and functioning of the algorithm and sometimes of the data it is trained on. There are also clear differences between the systems, which reflect different priorities of the two governments. New York's registry exclusively targets algorithms used by public agencies, facilitating greater citizen oversight of their government. The Chinese algorithm registry targets algorithms used by the private sector, giving citizens some information—and the state even more information—about them.

This priority reflects a dichotomy that underlies China's governance of data and AI. The regulations provide Chinese citizens with meaningful protection from Chinese companies, but they do not provide that same protection from the actions of the party-state. To many non-Chinese observers, this difference appears to be a fundamental contradiction, one that invalidates a core purpose of technology regulation. In China, this dichotomy is simply the reality.

Conclusion

Through the Policy Funnel

This paper has traced the policymaking process for two of China's foundational regulations on algorithms and AI, illustrating the diverse forces that spark and shape the country's technology policymaking. The initial spark for these regulations derived from a real-world change: a new technology application or business model that affected China's economic, social, and political order. In the case of the recommendation algorithm regulation, that change was the rise of algorithm-driven news apps like Toutiao and the new challenges they posed to content moderation and ideological control. Other real-world shifts that eventually found their way into the regulation included the rise of algorithm-driven food delivery apps and the adoption of algorithmic price discrimination by e-commerce platforms. In the case of deep synthesis, it was a technological innovation that enabled face-swapping applications that sparked a public outcry over their implications.

Those shifts were then filtered through the twin lenses of CCP ideology and, in some cases, Xi's own preferences. Algorithm-driven news apps undermined CCP conceptions of the role of media, and the party weighed in through a series of editorials. The impact of algorithms on Chinese migrant laborers similarly ran counter both to professed party doctrine and to Xi's push for "common prosperity," the latter being a factor contributing to the crackdown on China's leading technology platforms. While the rise of deepfakes did not go through a period of rigorous ideological analysis, the threats to information environments posed by fabricated images were clear to policymakers in China and around the world.

These issues then entered China's "world of ideas," where they were further defined and shaped by academic research, technical think tanks, investigative journalism, public opinion, and corporate lobbying. Here, the broad party-state imperative to do something about these technology applications was steadily molded into concrete interventions, such as new technical standards and certifications. Chinese sociologists and magazine journalists sharpened the focus on algorithmic exploitation, and a widespread public outcry brought this issue to the forefront. Meanwhile, corporate thought leadership reframed the debate on deepfakes, swapping the terminology and broadening the scope. All throughout this process, international discussions of these same topics permeated China's world of ideas, with China's state media and policy community consistently engaging with and sometimes adapting those ideas for the Chinese context.

All of these influences fed into bureaucratic attempts to formally regulate the technologies. These attempts were highly iterative, beginning with blunt corporate punishments before evolving into progressively more robust regulations. The CAC led the charge in both cases, relying on extensive consultations with Chinese legal academics and technologists to shape the text of the regulations. It also looped in related ministries and agencies, some of which added new provisions to address their own concerns, such as algorithm-fueled monopolistic behavior. What emerged from the end of this process were two of the world's earliest major regulations on algorithms and AI.

Generative AI and the Road Ahead

While the recommendation algorithm and deep synthesis regulations laid the foundations for China's AI governance regime, they were just the beginning. In November 2022, just five days after the deep synthesis regulation was signed, OpenAI shook up the world with the release of ChatGPT. The chatbot's ability to produce credible writing on innumerable topics wowed users around the world, including many in the CCP and China's government. The deep synthesis regulation technically already covered tools like ChatGPT—under the phrasing "technologies for generating or editing text content"—but it did not anticipate the power or popularity of this new generation of large language models.

Regulators quickly decided new rules were needed. Over the next eight months, the CAC and other regulators quickly produced a draft and then an "interim" generative AI regulation. Technological and economic shifts during this time were altering the state's calculus on trade-offs between political security and AI development, and the regulation was the subject of intense intellectual debate and bureaucratic wrangling. The next paper in this series will examine these debates and the continued evolution of China's AI governance regime, both at home and on the international stage.

Those developments followed a similar, though much accelerated, path through the policymaking process detailed in this paper. The process featured many of the same players—ministries, policy advisers, and technical organizations—as well as some new ones. And it further elucidated the key tensions and likely trajectories of China's AI regulations.

What comes next will hold profound implications for the development and governance of AI, both within China and around the globe. Over the next few years, Chinese companies will continue pushing AI products onto global markets, and Chinese diplomats will engage with and reshape international governance regimes. The country's domestic regulations will both constrain and enable those efforts, and understanding the individuals, ideas, and institutions guiding China is all the more important.

About the Author

Matt Sheehan is a fellow at the Carnegie Endowment for International Peace, where his research focuses on global technology issues, with a specialization in China's artificial intelligence ecosystem.

Acknowledgments

The author would like to thank Zhang Peng, Jamie Horsley, and Jeremy Daum for their helpful feedback and corrections to an early draft of this paper. The responsibility for the content and any remaining errors remain exclusively with the author.

Carnegie Endowment for International Peace

The Carnegie Endowment for International Peace is a unique global network of policy research centers around the world. Our mission, dating back more than a century, is to advance peace through analysis and development of fresh policy ideas and direct engagement and collaboration with decisionmakers in government, business, and civil society. Working together, our centers bring the inestimable benefit of multiple national viewpoints to bilateral, regional, and global issues.

Asia Program

The Asia Program in Washington studies disruptive security, governance, and technological risks that threaten peace, growth, and opportunity in the Asia-Pacific region, including a focus on China, Japan, and the Korean peninsula.

CARNEGIE
ENDOWMENT FOR
INTERNATIONAL PEACE

CarnegieEndowment.org