

# New Methods, Old Debates:

## *The Contested Formation of Computational Social Science in China*

Chengpang Lee<sup>1\*</sup>

Social Constellations  
1(2): 84–96  
DOI: 10.66630/sc.2026.  
0013  
<https://www.sc-world.org>

Received: May 21, 2026  
Revised: Jun 15, 2026  
Accepted: Jun 16, 2026

### Abstract

This paper explores the development of computational social science (CSS) in China via the model of contested field formation. This model highlights that Chinese CSS emerges amid tensions between institutionalization-driven structural forces and entrenched epistemological traditions, defying both the Western diffusion and the top-down dominance models. The model highlights that three external structural pressures fuel CSS's expansion in Chinese universities: the state's New Liberal Arts (*xinwenke*) policy, a central-government's higher education reform that pushes humanities and social sciences toward interdisciplinary, digital, and now AI+ transformation; a graduate labor market crisis that renders CSS training a career credential for humanities and social sciences students; and growing platform data and computational infrastructure enabling innovative research designs. These pressures face two internal constraints: a faculty shortage hindering effective New Liberal Arts implementation on the one hand and, on the other, the epistemological pushback from indigenization-oriented and autonomous-knowledge-system discussions that pivot on whether CSS can generate signature Chinese concepts or merely import universalist categories.

### Keywords

Computational Social Science; New Liberal Arts; Indigenization; Sociology of Knowledge

### Introduction

Computational social science (CSS)—defined as the development and application of computational methods to the scientific analysis of social phenomena (Lazer et al., 2009, 2020)—has emerged as a major paradigm in global social inquiry. From tracking information diffusion through social networks to modeling urban mobility and analyzing cultural production at scale, computational approaches

---

<sup>1</sup> School of Sociology, Central China Normal University, China

\***Corresponding Author.** [lzb1223@ccnu.edu.cn](mailto:lzb1223@ccnu.edu.cn)

have transformed the questions social scientists can ask and the evidence they can bring to bear (Salganik, 2018; Watts, 2007).

China presents a particularly illuminating case for studying CSS formation because computational methods are emerging at the same moment that Chinese social science is being asked to build an autonomous knowledge system of its own. The call for signature concepts drawn from Chinese experience is now one of the most important events shaping the field's intellectual horizon, yet its relationship to CSS has not been systematically discussed: can large-scale computational evidence help generate and test such concepts, or does CSS simply import categories whose Chinese fit remains unexamined?<sup>1</sup> No other major academic system has simultaneously deployed top-down state imperatives for computational transformation through the New Liberal Arts policy and the newer AI+ initiative associated with the 15th Five-Year Plan, confronted a severe employability crisis among humanities and social science graduates, maintained a rich tradition of sociological indigenization and qualitative fieldwork, and developed sizable CSS communities—all within the span of a decade. These conditions actively shape what Chinese CSS becomes.

This paper argues that CSS development in China is best understood as contested field formation. The field is pulled by structural forces demanding computational capacity and complicated by epistemological traditions contesting computation's adequacy for social reality. What emerges is neither smooth diffusion from West to East nor a product of state direction, but a field whose intellectual character is actively negotiated by actors with genuinely different visions of what CSS should be and what it can illuminate.

My analysis is organized around two connected questions. First: what structural forces drive CSS institutionalization in China, and how do the *xinwenke* (New Liberal Arts) initiative and the graduate labor market crisis create conditions in which CSS appears simultaneously as policy priority, employability signal, and curricular imperative universities cannot substantively fulfill? Second: what field-internal tensions shape what CSS becomes—specifically, how does the indigenization debate map onto the conflict between qualitative fieldwork traditions and computational approaches?

## Theorizing Computational Social Science

Foundational CSS literature emerged from Western institutional contexts. Lazer et al. (2009, 2020) established the field's intellectual foundations in Science, arguing that computational methods enable a paradigm shift in social inquiry. Conte et al. (2012) articulated a complementary European vision emphasizing social simulation and theory-method integration. Edelmann et al. (2020) provided a critical review acknowledging CSS's success in descriptive tasks alongside its struggles with explanatory and theoretical contribution—a limitation, this paper argues, that reflects structural and epistemological conditions as much as intellectual ones.

English-language studies of China's CSS remain limited. Zhou (2022) systematically documents rapid growth and key institutions. Wu et al.'s (2024) edited Routledge volume provides the most comprehensive English-language treatment to date. From a comparative standpoint, Milan and Treré's (2019) "Big Data from the South(s)" offers a framework for understanding how different institutional configurations produce different computational social sciences—a lens this paper applies to China with attention to the specific structural and epistemological forces at work.

Chinese-language scholarship offers a richer internal discourse than English-language audiences typically access. Wang Chengjun (2021) asks what a distinctively Chinese center of computational

---

<sup>1</sup> Within ethnological research, a large language model (LLM) have been developed to supply rigorously grounded, culturally authentic knowledge to students and scholars that study the development of the community for the Chinese nation.

knowledge production would look like, drawing on Latour's sociology of science. Meng et al. (2023) map social computing's opportunities and challenges with attention to the gap between methodological sophistication and theoretical development. Tang Shiping (2021) argues for CSS as a tool for scientific governance. The indigenization debate—Xie Yu's (2018) dismissal of indigenization as a pseudo-problem, Zhou Xiaohong's (2020) defense of it—provides the epistemological backdrop against which CSS advocates and qualitative researchers articulate their positions. Chen Jiajian (2025) articulates the epistemological claims of fieldwork traditions with precision and relevance to CSS's scope conditions in China's regulated information environment. The present paper integrates these sources to provide a framework existing literature lacks: an account of how structural forces driving CSS institutionalization interact with the epistemological traditions resisting it.

Existing reviews of Chinese CSS share a common limitation: they analyze the field from within the field's own self-understanding. For example, Zhou (2022) offers a descriptive inventory of growth trajectories, institutions, and phase transitions—the first of its kind in English, and invaluable as a baseline—but presents CSS as an academic success story without analyzing the structural forces (e.g., *xinwenke*, labor market) driving that growth or the epistemological resistances shaping its limits. Luo et al. (2021), valuable for methodological self-reflection, offers an insider perspective that brackets the policy conditions producing the methodological landscape it describes. Meng et al. (2023), written from a technocratic vantage, identifies capability gaps and opportunities without analyzing the structural or epistemological conditions generating them.

None of these reviews engages the *xinwenke* imperative, the graduate labor market crisis, the implementation deficit, or the autonomous knowledge system/indigenization debate as constitutive forces in CSS development—because their analytical stance is inside the field looking at its contents, rather than outside looking at its formation. The contested-formation framework occupies the latter position. It can see what the former cannot: that the most analytically significant questions about Chinese CSS concern not only what the field has produced but why it has taken the shape it has, and what forces determine which intellectual projects are possible within it.

Grounded in Bourdieu and Wacquant's (1992) field theory, this paper examines how fields take shape under overlapping external strain and internal disputes. It sketches the structural terrain of China's computational social science rather than undertaking exhaustive field analysis: new liberal arts mandates and labor market tensions counterbalance entrenched traditions of sociological indigenization and qualitative fieldwork, with no single force establishing hegemony.

The contested-formation framework complements rather than replaces structural analysis. State funding categories channel research toward certain topics; elite universities concentrate CSS capacity; platform corporations control data access. These structural facts are real and consequential. But they do not fully determine what Chinese CSS becomes intellectually. For that, we must attend to the epistemological contestation that structural analysis tends to bracket: genuine disputes about whether computation can adequately grasp Chinese social reality. Both structural and epistemological dimensions are necessary for a complete account.

## **CSS in China: Three Phases Reconsidered**

China's CSS exhibits three historical phases. Reading them through the contested-formation lens reveals not simply a story of growth and institutional consolidation, but a story of intensifying structural pressures and accumulating epistemological tensions—pressures and tensions that define the terrain on which confronting ideal scholars operate. The 2012 boundary reflects the convergence of platform maturation, state big-data policy, and expanding international collaboration; the 2019

boundary reflects the formalization of *xinwenke* and the regulatory tightening that reshaped data access conditions. Each transition redistributed conditions for CSS production while the epistemological contestations accumulated rather than resolved.

### *Phase One: Early Foundations (Pre-2012)*

During this phase, CSS imperative remained limited to a small number of pioneering scholars at elite universities and government institutes. Individual returnee scholars—Chinese researchers trained in the United States and Europe—introduced computational approaches at Tsinghua, Peking, and Fudan Universities. Their dominance was intellectual rather than institutional: they controlled methodological knowledge unavailable to most Chinese social scientists, and their international training provided cultural and linguistic capital that distinguished them within domestic academic hierarchies.

The Chinese Academy of Sciences' Institute of Sociology established exploratory working groups on computational approaches. Government interest in "social management" (*shehui guanli*) motivated early computational work on predicting and monitoring social dynamics, signaling the state's early recognition of computation's governance potential. Limited digital infrastructure and data constrained expansion. Critically, the indigenization debate had not yet crystallized in its modern form, and *xinwenke* as a policy framework did not exist. This phase represents CSS before its contested formation began in earnest: structural pressures were nascent, epistemological tensions were latent, and the confronting ideal scholars had not yet found the terrain they would later navigate.

### *Phase Two: Rapid Expansion (2012–2018)*

Phase Two witnessed dramatic growth and a decisive shift toward institutional production. The State Council's 2015 Big Data Development Action Outline and the 2017 AI Development Plan provided legitimacy and funding at strategic national scale. Platform corporations emerged as critical data gatekeepers: Weibo's national-scale social media data enabled public-opinion analysis at scope unachievable elsewhere; WeChat mapped communication networks; Alibaba and JD.com generated transaction records at unprecedented scale. Elite research institutions established CSS-relevant units—the Center for Social Computing and Parallel Management at the CAS Institute of Automation, CSS-affiliated faculty and programs distributed across Tsinghua's School of Social Sciences and Department of Computer Science, and Fudan's Center for Complex Decision Analysis.

The coordination between state, university, and platform was visible in concrete funding mechanisms. NSFC programs explicitly solicited big-data social science proposals; the 2017 implementation of China's Double First-Class initiative allocated funds for interdisciplinary programs at designated universities; platform corporations collaborated selectively with researchers on projects aligned with their commercial interests. International collaboration expanded as Chinese scholars participated in IC2S2 conferences and published internationally (Zhou, 2022). Simultaneously, the epistemological tensions surfaced with clarity: Xie Yu's (2018) provocation against indigenization as a pseudo-problem provoked responses from renown Chinese sociologists such as Zhou Xiaohong and Zhai Xuwei, mapping directly onto the emerging confrontation between CSS's universalist methods and the established qualitative-fieldwork tradition. Institutional expansion and epistemological contestation intensified in parallel without converging—the stage set for Phase Three's contested formation.

### *Phase Three: Maturation and Institutionalization (2019-Present)*

Phase Three is defined by two simultaneous processes that pull in different directions: the formalization of *xinwenke* as a central-government education policy and the tightening of data regulation. The 2019 Six Excellence One Summit Plan 2.0 and the 2020 *xinwenke jianshe xuanyan* (Declaration on the Construction of New Liberal Arts) established computational transformation as a policy task for humanities and social sciences; the more recent AI+ initiative associated with the 15th Five-Year Plan further extends this policy logic from digital literacy to AI-enabled teaching, research, and curriculum reform. Simultaneously, the 2021 Personal Information Protection Law (PIPL) and the 2021 Data Security Law imposed comprehensive restrictions on data collection, storage, and cross-border transfer, fundamentally reshaping the conditions under which computational research operates.

The conjunction reveals a structural irony: the state simultaneously imperated more CSS and regulated the data environment that makes CSS distinctive. Universities were advised to offer CSS education while their research pipelines were constrained by rules governing the platform data that had driven Phase Two's growth. Under the PIPL, large-scale collection of user-generated data requires explicit consent or a recognized lawful basis; cross-border data transfer triggers separate security assessments. Research designs routine in Phase Two now require platform partnerships or institutional reviews that concentrate feasibility at well-connected institutions.

The implementation deficit—the gap between *xinwenke* imperative and faculty capacity—became visible in this phase. Universities without qualified faculty found themselves administratively required to offer computational curricula they could not deliver substantively. The epistemological tensions, rather than resolving, deepened as CSS's institutional presence grew without engaging the methodological questions that its critics had been articulating since Phase Two. Phase Three is therefore not simply the maturation of an established field but the intensification of a contested formation—with structural pressures sharpened and epistemological resistances clarified.

### **Structural Forces: Imperative, Market, and the Implementation Gap**

Three structural forces drive CSS institutionalization in China from outside the scholarly community: a state imperative that places CSS at the center of higher education reform; a labor market crisis that makes computational skills an employability signal; and expanding platform data and computational infrastructure that make certain research designs newly possible. These forces converge to create intense institutional pressure for CSS expansion. But they do not, by themselves, produce coherent intellectual development—and the gap between the institutional pressure they create and the scholarly capacity available to respond to it is itself a defining feature of Chinese CSS's current character.

#### *The New Liberal Arts Initiative as State Imperative*

In November 2020, the Ministry of Education convened the New Liberal Arts Construction Working Conference at Shandong University, issuing the *xinwenke jianshe xuanyan* (Declaration on the Construction of New Liberal Arts). *Xinwenke* should be read here not as a disciplinary category but as a central-government education policy: it calls universities to reorganize humanities and social science education around national development, interdisciplinary training, digital technology, and the production of "interdisciplinary, cross-domain, composite talent" suited to socialist modernization

(Ministry of Education, 2020). The latest AI+ initiative in the 15th Five-Year Plan gives this policy a sharper technological form, making CSS a paradigmatic reform project rather than merely one research specialty among others.

This declaration formalized a direction established by the 2019 Six Excellence One Summit Plan 2.0—jointly launched by thirteen ministries—which committed to weaving digital technology into every domain of higher education, including disciplines with no prior tradition of quantitative or computational work. For CSS, the policy consequences were immediate. Universities that had hesitated to establish computational programs now had ministerial cover; faculty proposing new computational curricula could cite *xinwenke* requirements as justification; degree programs could be restructured under *xinwenke* labels that unlocked new funding streams and satisfied external review criteria.

The *xinwenke* framework also defines CSS's relationship to state goals in ways that shape its intellectual character. The declaration's emphasis on "Chinese characteristics" and "serving socialist modernization" signals that computational capacity is valued instrumentally—for governance applications and national competitiveness—rather than as an autonomous intellectual project. CSS that contributes to public-opinion analysis, smart governance, or social stability monitoring fits the *xinwenke* frame comfortably; CSS that investigates platform power, algorithmic accountability, or political inequality does not. The imperative creates institutional space for CSS while simultaneously defining that space's boundaries, shaping which projects receive institutional support and which are structurally awkward. Its breadth—encompassing activities from genuine computational research to adding digital tools to existing courses—allows institutions to satisfy *xinwenke* requirements at varying levels of substantive engagement, making formal program proliferation a poor indicator of actual CSS development.

### *Labor Market Pressure and the Employability Signal*

The *xinwenke* initiative did not emerge from purely academic considerations. Its policy context is inseparable from a structural crisis in the Chinese graduate labor market—and CSS's expansion is entangled with this crisis in ways that shape both who studies CSS and why institutions invest in it.

Chinese universities have graduated increasing numbers of humanities and social science students at precisely the moment when structural demand for such graduates has stagnated. Youth unemployment climbed sharply in 2023, with the official rate for 16–24 year-olds reaching 20.8% by May and 21.3% by June before the National Bureau of Statistics suspended publication of the series under the old methodology (National Bureau of Statistics of China, 2023). Graduate employment discussions consistently identify overqualification, major-job mismatch, and pressure to demonstrate applied skills as salient problems for nontechnical graduates. Policy documents acknowledge this indirectly: *xinwenke* declarations repeatedly invoke "composite talent" (*fuhexing rencai*), "application-type professionals" (*yingyongxing rencai*), and, more recently, AI+ integration under the 15th Five-Year Plan, recognizing that humanities and social science education should articulate a clearer labor-market rationale to justify the social resources it consumes.

Within this context, CSS carries an employability signal partly independent of its intellectual content. A humanities graduate who can demonstrate computational skills—programming, data visualization, text analysis, network methods—is positioned differently in the labor market than one who cannot. This signal operates even when training is superficial: the credential of having completed coursework labeled as data analysis communicates digital literacy to employers regardless of whether genuine research capacity was developed. Universities marketing CSS-inflected programs understand this, as do students navigating a difficult job market for whom the

computational credential is as much a defensive strategy as an intellectual choice. The growth of CSS programs, enrollments, and degree offerings in Phase Three may therefore reflect genuine intellectual transformation at elite institutions and employability branding at others—a distinction the implementation deficit makes visible.

### *The Implementation Deficit: Teacher Training and Capacity Gaps*

The most revealing structural tension in Phase Three is the implementation deficit: *xinwenke* imperatives interdisciplinary computational curricula, but most Chinese universities lack the qualified faculty to deliver them in ways that realize their stated intellectual aims.

Faculty development literature on *xinwenke* and CSS consistently identifies teacher capacity as a central bottleneck (Meng et al., 2023; Zhou, 2022). Existing humanities and social science faculty typically lack computational training adequate for teaching CSS; existing computer science faculty typically lack the social science grounding to connect computational tools to substantive research questions; scholars trained adequately in both domains are scarce and concentrated at a small number of elite institutions with international recruitment capacity. The same tension is visible in digital humanities. Traditional humanities fields are also adopting computational tools, but the direction and value of such work often remain uncertain: a network analysis of characters in *Dream of the Red Chamber*, for example, may demonstrate technical competence without necessarily showing how much it contributes to established *Honglouloumeng* scholarship. This produces a further anxiety among humanities scholars that AI+ reform may train them to become second-rate programmers rather than better interpreters of culture, texts, and social life. Universities that could previously staff a social statistics course with a quantitatively trained sociologist cannot staff a CSS course with a scholar who knows both computational methods and fieldwork traditions—because such scholars are rare in China's academic labor market, where computational and humanistic training have historically been separated by disciplinary structures.

The consequence is institutional divergence rather than the convergence that uniform *xinwenke* imperatives might suggest (Wu et al., 2024; Zhou, 2022). At Tsinghua, Peking, or Fudan, institutions that can recruit internationally trained scholars and offer competitive packages to computational social scientists—CSS programs reflect genuine integration of methodological and substantive training. At provincial universities unable to compete for such faculty, CSS often means adding data-visualization workshops to existing methods courses, or introducing programming modules whose connection to social science research questions is left to students to discover. The imperative is formally universal; the capacity to meet it substantively is steeply hierarchical.

This concentration reproduces institutional hierarchy through a mechanism different from political control but equally consequential: the uneven distribution of teaching capacity across an educational system where *xinwenke* creates identical formal requirements for all universities but the human capital to meet those requirements is available to very few. The confronting ideal scholars—whose intellectual projects navigate the field's tensions most productively—are overwhelmingly located at elite institutions, not because CSS requires elite credentials but because only elite institutions have assembled the conditions in which a scholar can genuinely engage the tensions between structural imperative and epistemological resistance rather than simply satisfying the former while ignoring the latter.

## Field-Internal Tensions: Indigenization and the Qualitative Challenge

The structural forces described in Section 3 create institutional pressure for CSS from outside the scholarly community. They encounter, however, two forms of resistance from within Chinese social science: an epistemological challenge from scholars committed to sociological indigenization and autonomous knowledge construction, and a methodological challenge from scholars committed to qualitative fieldwork. Understanding these resistances is essential for understanding what CSS becomes in China, because the most intellectually significant CSS emerges not from scholars who ignore these challenges but from those who engage them.

### *The Indigenization Debate as Epistemological Backdrop*

In 2018, Princeton and Peking University sociologist Yu Xie published a provocation in *Sociological Studies* arguing that indigenization—the project of developing distinctively Chinese social science rooted in Chinese cultural particularity—is a pseudo-problem (Xie, 2018).<sup>2</sup> All social science, Xie contended, must ultimately conform to evidence-based scientific standards. What presents itself as distinctively Chinese sociology more often reflects methodological immaturity or political defensiveness than genuine cultural authenticity. Rigorous, evidence-based inquiry—whether quantitative, computational, or qualitative—is not culturally particular but universally applicable across contexts.

Zhou Xiaohong and Zhai Xuewei responded forcefully. For Zhou (2020), whether Western social science categories—developed in European liberal-individualist contexts—apply without modification to Chinese social arrangements is a genuine epistemological question. Chinese concepts of relational selfhood, face (*mianzi*), reciprocal obligation (*guanxi*), and collective memory shaped by socialist history may require different analytical frameworks than those developed for Anglophone individualist societies. To apply Western frameworks without examining their assumptions is not methodological universalism but methodological imperialism in the garb of rigor. Zhai (2018) pushed further: Chinese social science needs concepts adequate to Chinese social reality rather than imported instruments whose fit is assumed rather than demonstrated.

This debate maps directly onto CSS's intellectual situation in China. CSS's universalist episteme—algorithms generalize; methods should replicate across contexts; evidence-based standards transcend cultural particularity—aligns with Xie's position. From this perspective, CSS is the methodological fulfillment of the argument against indigenization: rigorous, scalable, internationally commensurable, and unconcerned with cultural specificity as a methodological obstacle. If computational methods can be validated on Chinese data and produce replicable findings that meet international scholarly standards, that is sufficient methodological legitimacy.

But the indigenization critics identify a deeper problem that methodological sophistication does not resolve. If the concepts organizing CSS research—public opinion, polarization, social capital, information cascade—presuppose institutional and cultural configurations structured differently in China, then computational rigor applied to these concepts risks mistaking precision for accuracy. The prior question of whether Chinese online discourse is the right object to which the concept of public opinion applies is a conceptual question that no methodological choice resolves; answering it requires theoretical work that CSS tools do not themselves provide. This is a genuine challenge to

<sup>2</sup>The indigenization debate's key texts include Xie (2018), who discusses misunderstandings in Chinese sociology's indigenization debate, and Zhou (2020), who examines whether indigenization should be understood as a narrow or broad, or as a pseudo-problem or true reality. Zhai Xuewei's position is developed across multiple publications in *Chinese Journal of Sociology and Sociological Studies* between 2018 and 2022.

CSS's intellectual foundations that the field's most serious Chinese practitioners have begun to engage—and that the confronting ideal scholars take most seriously.

### *Qualitative Fieldwork and the Limits of Computation*

Alongside the indigenization debate, Chinese sociology maintains a strong tradition of qualitative fieldwork—rural ethnography, community studies, life-history interviews, institutional analysis—with roots in Fei Xiaotong's foundational studies of Chinese rural society and reconstructed through the reform era's engagement with rapid social transformation. This tradition does not merely constitute conservative resistance to computational newcomers; it represents a substantive epistemological position with its own analytical strengths that are not reducible to pre-digital nostalgia.

Chen Jiajian's (2025) analysis of qualitative research in comparative methodology articulates the position with precision relevant to CSS's scope conditions: qualitative research's distinctive contribution lies in its naturalness—the capacity to obtain data within research subjects' actual working and living situations rather than in controlled or observational conditions that alter what is visible—its openness to unexpected findings that the researcher did not anticipate, and its ability to situate observations in the contexts that make meaning possible (Chen, 2025). A computational analysis of migration patterns may capture scale, distribution, and temporal dynamics with impressive precision; it does not, by itself, explain why specific migrants make specific choices in specific local contexts shaped by specific institutional histories, relationships, and negotiations—which is often the sociologically interesting question.

This creates productive tension rather than simple incompatibility. Scholars who integrate computational and qualitative approaches—using computational analysis to identify patterns and then deploying fieldwork to interpret them—produce accounts that neither method alone can achieve. King et al. (2013) demonstrate that rigorous computational research on Chinese political institutions is both feasible and internationally impactful; their work on censorship selection logic combines computational pattern detection with theoretical interpretation of political logics that drive the patterns. Neither component alone produces the insight that their combination achieves.

The qualitative tradition also raises questions about CSS's scope conditions particularly acute in China's regulated information environment. Methods designed to analyze public social media discourse capture what users are willing to say in semi-public contexts under platform monitoring and content regulation; they do not capture negotiated meanings of everyday interaction, tacit knowledge governing social practice, or the gap between public expression and private belief. Where politically sensitive expression is self-censored or coded in ways automated analysis misidentifies, the gap between social media data and social reality may be systematically larger than in less-restricted environments. Chinese qualitative sociologists who point to this gap are not defending their positions against competition; they are identifying a genuine limitation of CSS's primary empirical materials in the Chinese context. Engaging this critique is part of what distinguishes the confronting ideal scholars.

### *A Path from Qualitative and Historical to Computational Social Science*

In my own research and experience, I first encountered Computational Social Science (CSS) while working on my dissertation back in 2015. The institution where I conducted my research houses a wealth of rich historical materials. Traditionally, the methodological approaches of historical sociologists are similar to those of historians. Beyond using software such as MaxQda to assist with

the coding process, I considered using Python to extract information that MaxQda could not provide—for instance, calculating word frequencies across an entire corpus. I wrote my first Python script to extract data from each volume, which I used in my dissertation and later in my monograph (Lee, 2025).

While continuing my work in the traditional qualitative research field, I developed an interest in CSS methodologies and the new forms of knowledge production they bring to social science research. To further my learning, I began formal academic training and field inquiry by enrolling in the online Computer Science degree programme offered by the University of London in 2021, with a concentration in Data Science. In this short span of time, the AI revolution has triggered a new paradigm shift within CSS. I started teaching AI and Society and Computational Social Science in my new academic position in 2026. The tools I now teach my students are vastly different from those I learned back in 2015. AI agents have taken over the tedious coding and web scraping tasks that once posed major obstacles for me. Some datasets I was unable to analyse at the time due to time constraints are now readily analysable with the support of AI agents. By the time this paper is finalized, 2026 has been widely hailed as the inaugural year of AI agents.

### *Research Themes and What the Tensions Foreclose*

The intersection of structural pressures and internal tensions shapes not only how CSS develops institutionally but what research it produces. Research themes in Chinese CSS reflect the interests and constraints of both the dominant structural forces and the field-internal tensions (King et al., 2013; Wu et al., 2024; Zhou, 2022). Public opinion and political communication dominate: how attitudes form on Weibo, how information diffuses under regulatory constraints, and how sentiment relates to policy are questions that satisfy state priorities, platform commercial interests, and CSS methodological toolkits simultaneously. Social network analysis builds on sociological traditions of studying *guanxi*. Urban studies apply computational methods to mobility data from mobile phones and ride-sharing platforms. Economic sociology examines e-commerce behavior at unprecedented scale.

Attitudes toward computational social science (CSS) vary substantially across different social science disciplines, although the pattern should be read as an indicative disciplinary contrast rather than a comprehensive survey. In journalism and communication studies, text-as-data and sentiment analysis have become especially visible research strategies (Grimmer & Stewart, 2013; Wu et al., 2024). Economics and policy-oriented computational work have more readily incorporated simulation, prediction, and AI-assisted modelling (Meng et al., 2023). In contrast, CSS remains a more marginal and contested influence within sociology. This disciplinary discrepancy can be partially explained by the theoretical and methodological tensions discussed earlier.

Contemporary Chinese CSS scholarship notably omits several dominant research themes prevalent in Western computational social science, including algorithmic bias detection, political polarization, critical platform power critique, and digital inequality studies. Such research lacunae stem not from isolated factors, but from the intertwined interplay of structural constraints and epistemological tensions within the field.

State research funding frameworks steer scholarly inquiry toward governance-aligned research topics, while stringent data regulatory policies limit academic access to proprietary platform data—an essential resource for rigorous algorithmic auditing. Compounding these structural barriers, the longstanding qualitative research tradition in sociology raises pointed critiques regarding the contextual limitations of CSS methodologies, particularly for sensitive research domains where pervasive self-censorship systematically distorts computational empirical data. Furthermore,

scholars pursuing critical inquiries into sensitive societal issues often opt for overseas academic appointments, situating their critical CSS research outside the scope of domestic academic output.

Ultimately, accounting for these persistent research omissions in Chinese CSS necessitates a holistic analysis that integrates both structural institutional constraints and ongoing epistemological contestations shaping the field's research boundaries.

## **Conclusion: Toward a Contested-Formation Account**

This paper has traced CSS development in China through three historical phases, showing how structural pressures intensified and epistemological tensions accumulated in each phase, culminating in Phase Three's contested formation: *xinwenke* initiative that outruns teacher capacity; labor market signals that reduce CSS to employability branding in many institutions; an indigenization debate that contests computation's adequacy for Chinese social reality; a qualitative tradition that identifies genuine scope conditions for what CSS's primary materials can reveal and a rooted suspicion toward the epistemological foundation of CSS.

The contested-formation framework offers advantages over both the diffusion model and the dominance model for understanding CSS development. The diffusion model—CSS arrived from the West, spread through Chinese universities, and adapted to local conditions—misses the structural forces that drove institutionalization (*xinwenke*, labor market) and the internal resistances that shaped what was institutionalized. The dominance model—the state, elite universities, and scholar-brokers control the field—identifies real structural concentrations but brackets the epistemological contestation that gives CSS its intellectual character. The contested-formation framework incorporates both structural power and intellectual contestation, showing how they interact to produce a field whose development cannot be read off from either set of forces alone. It also identifies the analytically significant actors more precisely: not the most powerful institutions but the scholars who engage the tensions most productively.

The paper's core theoretical proposition, offered as an invitation to comparative analysis: CSS formation, wherever it occurs, is shaped by the intersection of external structural pressures and internal epistemological traditions. Where both are strong and neither has achieved dominance, the result is a contested field whose intellectual character reflects ongoing negotiation rather than institutional settlement. Understanding CSS development in any national context requires attending to both dimensions—to what structural forces are pushing CSS institutionalization and to what disciplinary traditions are resisting, qualifying, or attempting to redirect computational approaches.

Understanding China's CSS development matters for global social science. The conditions of CSS formation vary substantially across political economies—that variation is not deviation from a Western norm but a global range of genuinely different formations, each with distinctive capabilities and constraints. The structural pressures visible in China—state imperatives for computational education, graduate labor market distress, teacher training deficits—are not uniquely Chinese; versions are visible in other emerging CSS communities. The epistemological tensions—between universalist computation and local knowledge traditions, between scalable methods and contextually dense fieldwork—are globally distributed, though they take particularly sharp form in China. China is therefore not simply an exceptional case or a representative case. It is a high-intensity case in which party-state policy capacity, platform infrastructure, labor-market pressure, and an unusually self-conscious indigenization debate converge. Comparing China with other CSS formations can clarify which elements travel across contexts and which depend on China's distinctive institutional configuration.

## Declaration of Conflicting Interests

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

## Funding

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

## ORCID

Chengpang Lee <https://orcid.org/0000-0003-1240-7388>

## References

- Bourdieu, P., & Wacquant, L. J. D. (1992). *An invitation to reflexive sociology*. University of Chicago Press.
- Chen, J. (2025). Qualitative research in comparative method: Tradition, issues, and development trends [比较方法中的质性研究：传统、问题与发展趋势]. *Academic Research*, (4).
- Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertész, J., Loreto, V., Moat, S., Nadal, J.-P., Sánchez, A., Nowak, A., Flache, A., San Miguel, M., & Helbing, D. (2012). Manifesto of computational social science. *The European Physical Journal Special Topics*, 214, 325–346. <https://doi.org/10.1140/epjst/e2012-01697-8>
- Edelmann, A., Wolff, T., Montagne, D., & Bail, C. A. (2020). Computational social science and sociology. *Annual Review of Sociology*, 46(1), 61–81. <https://doi.org/10.1146/annurev-soc-121919-054621>
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3), 267–297. <https://doi.org/10.1093/pan/mps028>
- King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, 107(2), 326–343. <https://doi.org/10.1017/S0003055413000014>
- Lazer, D. M. J., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N. A., Contractor, N., Fowler, J. H., Gutmann, M., Jebara, T., King, G., Macy, M. W., Roy, D., & Van Alstyne, M. (2009). Computational social science. *Science*, 323(5915), 721–723. <https://doi.org/10.1126/science.1169410>
- Lazer, D. M. J., Pentland, A., Watts, D. J., Aral, S., Athey, S., Contractor, N. S., Freelon, D., Gonzalez-Bailon, S., King, G., Margetts, H., Nelson, A., Salganik, M. J., Strohmaier, M., Vespignani, A., & Wagner, C. (2020). Computational social science: Obstacles and opportunities. *Science*, 369(6507), 1060–1062.
- Lee, C. (2025). *The rise of Tzu Chi: The making of a global Buddhist movement*. University of British Columbia Press.

- Luo, J., Chen, H., & Wang, F. (2021). Computational social science in the era of big data: Developments and reflections [计算社会科学在大数据时代的发展与反思]. *Sociological Studies*, 36(4), 1–25.
- Meng, T., Huang, K., Liang, Y., Qi, J., Lü, X., & Zhang, H. (2023). Social computing and social intelligence: Opportunities and challenges [社会计算与社会智能：机遇与挑战]. *Bulletin of the National Natural Science Foundation of China*, 37(3), 439–450.
- Milan, S., & Treré, E. (2019). Big data from the South(s): Beyond data universalism. *Television & New Media*, 20(4), 319–35. <https://doi.org/10.1177/1527476419837739>
- Ministry of Education of the People's Republic of China. (2020). Declaration on the construction of new liberal arts [新文科建设宣言].
- National Bureau of Statistics of China. (2023). Surveyed unemployment rate in urban areas, May–June 2023.
- Salganik, M. J. (2018). *Bit by bit: Social research in the digital age*. Princeton University Press.
- Tang, S. (2021). Computational social science and the future of scientific decision-making [计算社会科学与科学决策的未来]. Fudan University Center for Complex Decision Analysis Working Paper.
- Wang, C. (2021). Rethinking the logic of computational social science: Based on Latour's concept of the center of calculation [重新思考计算社会科学的逻辑：基于拉图尔“计算中心”概念]. *Nanjing Social Sciences*, (4).
- Watts, D. J. (2007). A twenty-first century science. *Nature*, 445, 489.
- Wu, X., Zhang, Y., & Cai, T. (Eds.). (2024). *Computational social science: Application in China studies*. Routledge.
- Xie, Y. (2018). Overcoming misunderstandings in Chinese sociology's indigenization debate [走出中国社会学本土化讨论的误区]. *Sociological Studies*, 33(6), 1–13.
- Zhai, X. (2018). Is the indigenization of sociology a pseudo-problem: A response to Xie Yu [社会学本土化是个伪问题吗：与谢宇商榷]. *Exploration and Free Discussion*, (10).
- Zhou, X. (2020). Indigenization of sociology: Narrow or broad, pseudo-problem or true reality [社会学本土化：狭义还是广义，伪问题还是真问题]. *Sociological Studies*, 35(1), 1–24.
- Zhou, Y. (2022). Computational social sciences in China: A decade of development [Preprint]. SocArXiv. <https://doi.org/10.31235/osf.io/wktz9>