

**IN SEARCH OF ACTIVIST LAWYERS IN CHINA:
A TIME MACHINE BACK TO THE HEYDAY OF
POLITICAL ACTIVISM IN THE CHINESE LEGAL
PROFESSION**

ETHAN MICHELSON*

ABSTRACT

An analysis of a large corpus of online discussions—posted mostly in 2004—amongst members of the largest online community of lawyers that ever existed in China supports much of the existing literature about Chinese lawyers’ politically liberal values. At the same time, it reveals that Chinese lawyers’ political discourse was far from monolithic. By bringing into high relief the plurality of Chinese lawyers’ voices, this article contributes an unprecedented effort to disaggregate their worldviews. Previous research suggests that lawyers’ online discourse was an important catalyst of politically liberal activism in China. Findings reported in this article reveal for the first time that Chinese lawyers also called for mobilization in support of pro-regime, illiberal, and nationalistic causes. Finally, at precisely the moment of the alleged heyday of China’s so-called “*weiquan* movement,” there was scarcely a trace of political activism under the mantle of *weiquan*, or “rights defense,” in the Chinese legal profession.

Introduction.....	351
I. The Prevailing Portrait of Lawyers-Cum-Activists in China	353
II. Data and Methods	357
A. Bulk Downloading	358
B. Building the Archive: A Relational Database	360
C. Topic Modelling.....	367

* Ethan Michelson, Associate Professor of Sociology and Law, Indiana University Bloomington. I presented preliminary findings of this research at the 2018 Annual Meeting of the Law & Society Association, Toronto, Canada, June 7-10, and a draft of this article at the 2019 Wisconsin International Law Journal Annual Symposium, Madison, WI, April 5. I owe a special thanks to Sida Liu for initially educating me about the All China Lawyers Association’s Forum in 2004, to Kathryn Hendley for giving me an incentive fifteen years afterwards to write this article about it, and to Carole Silver for valuable feedback. I am also grateful for the assistance of Elizabeth Astrup, Zhaodi Chen, Andrew Hill, and Zuoyu Tian. All remaining flaws are my own.

III. Plan of Analysis.....	369
IV. Findings	372
A. Step One.....	372
B. Step Two	377
C. Step Three	380
D. Step Four.....	388
1. General Interest: Topic 1.	389
2. Lawyers' Plight: Topic 2.	390
3. Political Liberalism: Topic 3.	391
4. Nationalistic: Topic 5.....	394
5. Forum-Related: Topic 6.....	397
6. Professional Regulation: Topic 15.....	397
7. Law Firms: Topic 17.....	398
V. Discussion and Conclusions	400
Appendix.....	402

INTRODUCTION

In 2004, Chinese lawyers had a lot on their minds. Their lives and careers were characterized by acute growing pains punctuated by an array of professional vulnerabilities, including: deadbeat clients who absconded with lawyer fees; predatory law firms exacting exorbitant rents from their lawyers; unauthorized competition from unlicensed practitioners; Janus-faced bar associations that seemed more interested in collecting fees from lawyers than in protecting their interests; and hostile state actors obstructing and intimidating lawyers trying to do their work.¹ At this point licensed lawyers numbered about 145,000, a 60% increase over the previous decade and an astonishing 600% increase over the previous two decades.² Since then their numbers have almost tripled yet again.³ By the early 2000s, Chinese lawyers had been jettisoned *en masse* from the state sector and had completed the transition to an almost fully private profession, which paradoxically served to aggravate their poorly

¹ Ethan Michelson & Sida Liu, *What Do Chinese Lawyers Want? Political Values and Legal Practice*, in CHINA'S EMERGING MIDDLE CLASS: BEYOND ECONOMIC TRANSFORMATION 310 (Cheng Li ed., 2010).

² Ethan Michelson, *Lawyers, Political Embeddedness, and Institutional Continuity in China's Transition from Socialism*, 113 AM. J. SOC. 352, 368 (2007); *Chinese Yearbook of Lawyers 2004* (中国律师年鉴2004) 256 (2005), available at <https://perma.cc/7NR2-VBZM>.

³ NAT'L BUREAU OF STATISTICS OF CHINA, CHINA STATISTICAL YEARBOOK (中国统计年鉴), <http://www.stats.gov.cn/tjsj/ndsj/2019/indexeh.htm> (last visited Feb. 19, 2020), available at <https://perma.cc/UV9P-ZU7B>.

recognized and weakly protected professional status and interests vis-à-vis the government and the public alike. These conditions helped radicalize their political values and galvanize their political activism.⁴

Chinese lawyers' anxieties, concerns, and hopes are preserved in a unique archive I created in 2005 of their online discussions hosted by the All China Lawyers Association ("ACLA") on its interactive forum (中国律师网互动社区). Although their online discussions were not entirely unconstrained by censorship (including self-censorship),⁵ ACLA forum participants were notably free—particularly by today's standards—to post messages that would be unlikely to survive more than a few hours in today's China.⁶ The ACLA forum was the first and last nationwide community hub for lawyers. Its participants were likely a non-trivial portion of the total population of lawyers at the time. After it went offline in 2008, Chinese lawyers have built and maintained social networks in other ways. However, their online networking behavior on newer social media platforms (e.g., Weibo, WeChat, etc.) is both more fragmented and subject to more intrusive official scrutiny and censorship.⁷ In short, my archive is a unique and valuable time machine back to what scholars have portrayed as the heyday of rights activism and political activism in the Chinese legal profession.

I am not the first to analyze online discussions from the ACLA forum. Sida Liu and Terence Halliday have extensively analyzed the contents of sixty discussion threads from this website to support their arguments about Chinese lawyers' yearnings for political liberalism.⁸ I am, however, the first to situate the specific topic of interest to Liu and Halliday in the context of topics salient to a broader cross-section of Chinese lawyers in 2004. This exercise is both a theoretical opportunity and a methodological challenge. My archive of lawyers' online

⁴ See generally Michelson & Liu, *supra* note 1; Michelson, *supra* note 2; SIDA LIU & TERENCE C. HALLIDAY, *CRIMINAL DEFENSE IN CHINA: THE POLITICS OF LAWYERS AT WORK* (2016); Hualing Fu & Richard Cullen, *Climbing the Weiquan Ladder: A Radicalizing Process for Rights-Protection Lawyers*, 205 CHINA Q. 40 (2011).

⁵ Terence C. Halliday & Sida Liu, *Birth of a Liberal Moment? Looking through a One-Way Mirror at Lawyers' Defense of Criminal Defendants in China*, in *FIGHTING FOR POLITICAL FREEDOM: COMPARATIVE STUDIES OF THE LEGAL COMPLEX AND POLITICAL LIBERALISM* 65, 77 (Terence C. Halliday et al. eds., 2007).

⁶ LIU & HALLIDAY, *supra* note 4, at 151–53; see generally Gary King et al., *How Censorship in China Allows Government Criticism but Silences Collective Expression*, 107 AM. POL. SCI. REV. 326 (2013); MARGARET E. ROBERTS, *CENSORED: DISTRACTION AND DIVERSION INSIDE CHINA'S GREAT FIREWALL* (2018).

⁷ See generally King et al., *supra* note 6.

⁸ See generally LIU & HALLIDAY, *supra* note 4.

discussions is a portal to the past through which we can travel back in time to revisit the issue of lawyers' political values and aspirations. At the same time, reconstructing over 20,000 conversations held by over 8,000 participants poses unique methodological challenges that only now, fifteen years after I started this project, have I surmounted with at least some measure of success. When I created the archive in 2005, appropriate analytical tools and requisite computing power were both far beyond my reach. Technological and methodological advancements have finally caught up with the data.

In our time travel back to 2004, we will find empirical support for previous scholarly claims about politically liberal values and activism in the Chinese bar. We will also find a more nuanced and complex landscape of discussion topics beyond the politically liberal discourse that thus far has dominated the attention of scholars. The most salient topics of discussion on the ACLA forum have little to do with lawyering at all. Topics with political overtones are only a subset of all topics related to the legal profession, most of which concern substantive legal matters such as contracts, commercial law, labor disputes, and so on (some of which, of course, relate to politics). Conversations that were squarely animated by politics were located across the ideological spectrum. Chinese lawyers spoke not only the language of political liberalism; they also spoke the language of nationalism. They not only criticized the Chinese state and embraced democracy and rule-of-law norms concerning judicial independence and the separation of powers; they were also staunchly pro-regime and bitterly critical of—and openly hostile towards—the United States and Japan. They not only lamented endemic corruption in general and judicial corruption in particular, which they attributed to the absence of rule of law; they also advocated Maoist-style political purges—the very antithesis of due process—as a means of combatting corruption. They not only called for mobilization in pursuit of political liberalism; they also called for the military invasion of Taiwan and for violent action against the people and territorial claims of Japan.

I. THE PREVAILING PORTRAIT OF LAWYERS-CUM-ACTIVISTS IN CHINA

Our knowledge about politically liberal activist lawyers is vastly disproportionate to their prevalence and influence in China. Indeed, much of what we know about Chinese lawyers as a whole concerns their liberally-inspired political activism. To be sure, even if they account for

“fewer than one in a thousand licensed lawyers,”⁹ there are important reasons to study “the fringe” of the Chinese legal profession.¹⁰ There are also important reasons to contextualize the fringe and to consider the larger whole—and the diversity of its composition—of which they are a part. Pursuing this latter task is one of my goals in this article.

Beginning in the mid-2000s, journalists and scholars heralded the arrival of a new and distinct species of lawyer in China, the so-called *weiquan* (维权) lawyer. The earliest English-language academic publications (I could find) containing this term appeared in 2006.¹¹ The first English-language newspaper article (I could find) containing this term appeared the same year in *The New York Times*,¹² two years before *The China Journal* published Fu Hualing and Richard Cullen’s classic article containing the term in its title.¹³ Even if they do not contain this specific term, earlier media reports about the arrest and prosecution of Chen Guangcheng (陈光诚) refer to him and his prominent activist lawyer supporters, including Gao Zhisheng (高智晟), Teng Biao (滕彪), and Xu Zhiyong (许志永), as “human rights lawyers.”¹⁴

According to Fu and Cullen, rights activists, including self-styled *weiquan* lawyers, were emboldened by—and arose out of—a flurry of lawmaking in 1989, 1993, 1995, 1996, 1997, 2003, and 2004 intended (at least symbolically if not practically) to provide checks against arbitrary state power, to advance the “rule of law,” and to equip citizens with legal tools for asserting their “human rights.”¹⁵ Important legal victories won by activist lawyers include the 2001 Qi Yuling (齐玉苓) case, the 2003 Sun Zhigang (孙志刚) case, and the 2004 Zhang Xianzhu (张先著) case, all of which, according to the existing literature, were milestones in and helped

⁹ EVA PILS, CHINA’S HUMAN RIGHTS LAWYERS: ADVOCACY AND RESISTANCE 1 (2015).

¹⁰ Rachel E. Stern, *Activist Lawyers in Post-Tiananmen China*, 42 L. & SOC. INQUIRY 234, 242 (2017).

¹¹ See generally Eva Pils, *Asking the Tiger for His Skin: Rights Activism in China*, 30 FORDHAM INT’L. L.J. 1209 (2006); Keith J. Hand, *Using Law for a Righteous Purpose: The Sun Zhigang Incident and Evolving Forms of Citizen Action in the People’s Republic of China*, 45 COLUM. J. TRANSNAT’L L. 114 (2006).

¹² Joseph Kahn, *Rights Group Urges China to End Curbs on Lawyers*, N.Y. TIMES, Dec. 11, 2006, at A6.

¹³ Hualing Fu & Richard Cullen, *Weiquan (Rights Protection) Lawyering in an Authoritarian State: Building a Culture of Public-Interest Lawyering*, 59 CHINA J. 111 (2008).

¹⁴ Maureen Fan, *Blind Chinese Activist Gets 4 Years Supporters See Sentence as Punishment for Exposing Forced Abortions, Sterilizations*, WASH. POST, Aug. 25, 2006, at A09; see also Jim Yardley, *China Tries Rights Lawyer, Barring His Kin and Counsel*, N.Y. TIMES, Dec. 14, 2006, at A14.

¹⁵ Fu & Cullen, *supra* note 13, at 123.

fuel the rise of the so-called *weiquan* movement.¹⁶ Chen Guangcheng's well-publicized legal activism spanned 1997 and 2005.¹⁷ In short, scholarly accounts place the dawn and heyday of *weiquan* activism in the early 2000s. For this reason, the degree to which *weiquan* discourse is conspicuous in my archive of ACLA forum discussions should reflect the degree to which it was relevant and important in the Chinese legal profession as a whole.

The term "*weiquan*," as an adjective modifying the word "lawyer," has been variously translated as "right-protection," "rights-defense," "rights activist," "human rights," "civil rights," "cause," and "public interest."¹⁸ Members of this species are often simply called "rights lawyers" or "rights defenders," and are allegedly part of a larger "rights-defense movement."¹⁹ A key defining characteristic of *weiquan* lawyers includes their use of law to advance larger causes, such as the protection of human rights, beyond the immediate case at hand.²⁰ Lawyers who do not adopt the *weiquan* moniker may nonetheless be activists "committed to the ideals of political liberalism and actively fight against the Party-state."²¹ As adherents to political liberalism, they attach importance to a moderate state, the rule of law, basic legal freedoms, and civil society.²²

A dizzying array of typologies classify Chinese lawyers according to the nature and degree of their activism. *Weiquan* lawyers are subdivided into "moderate," "critical," and "radical" categories.²³ Elsewhere is a more comprehensive typology of "progressive elites" and "political activists"—

¹⁶ See generally Thomas E. Kellogg, *Constitutionalism with Chinese Characteristics? Constitutional Development and Civil Litigation in China*, 7 INT'L. J. CONST. L. 215 (2009); Hand, *supra* note 11; Pils, *supra* note 11; Fu & Cullen, *supra* note 4, at 42; PILS, *supra* note 9, at 48–49.

¹⁷ See CHEN GUANGCHENG, *THE BAREFOOT LAWYER: A BLIND MAN'S FIGHT FOR JUSTICE AND FREEDOM IN CHINA* (2015).

¹⁸ See generally Fu & Cullen, *supra* note 13; Fu & Cullen, *supra* note 4; Hualing Fu, *Human Rights Lawyering in Chinese Courtrooms*, 2 CHINESE J. COMP. L. 270 (2014); Hualing Fu, *The July 9th (709) Crackdown on Human Rights Lawyers: Legal Advocacy in an Authoritarian State*, 27 J. CONTEMP. CHINA 554 (2018); Eva Pils, 'Disappearing' China's Human Rights Lawyers, in *COMPARATIVE PERSPECTIVES ON CRIMINAL JUSTICE IN CHINA* 411 (Mike McConville & Eva Pils eds., 2013); PILS, *supra* note 9; Pils, *supra* note 11; CHEN, *supra* note 17, at 131.

¹⁹ See generally MARY E. GALLAGHER, *AUTHORITARIAN LEGALITY IN CHINA: LAW WORKERS, AND THE STATE* (2017); Chongyi Feng, *The Rights Defence Movement, Rights Defence Lawyers and Prospects for Constitutional Democracy in China*, 1 COSMOPOLITAN CIV. SOCIETIES J. 150 (2009); PILS, *supra* note 9; Ching Kwan Lee, *Rights Activism in China*, 7 CONTEXTS 14 (2008).

²⁰ See generally Fu & Cullen, *supra* note 13; PILS, *supra* note 9.

²¹ LIU & HALLIDAY, *supra* note 4, at 90; see Michelson & Liu, *supra* note 1, at 323 (reporting from a large nationwide survey that the political values of self-identified *weiquan* lawyers were no different from those of lawyers who did not identify as such).

²² Michelson & Liu, *supra* note 1.

²³ See generally Fu & Cullen, *supra* note 13.

who are politically liberal—and “pragmatic brokers” and “routine practitioners”—who are not politically liberal. Political activists are further disaggregated into “notable activists” and “grassroots activists.”²⁴

A similarly comprehensive typology consists of conservative “rights reactionaries,” “indifferent lawyers,” “informed pragmatists,” and politically liberal “rights activists.”²⁵ New categories of politically liberal activist lawyers have emerged, including the “die-hard lawyer” (死磕派律师)²⁶ and the “progressive proceduralist.”²⁷ As ideal types, they were developed largely for heuristic purposes in the absence of any empirical assessment of the relative size and importance of each category. As oxymoronic as it may seem, some *weiquan* lawyers are not even licensed lawyers,²⁸ and some *weiquan* lawyers are motivated as much by fee potential as by a larger social or political cause.²⁹

Insofar as activist lawyers have figured prominently, if not exclusively, in most efforts to classify Chinese lawyers, few attempts have been made to disaggregate the profession as a whole and thus to situate the politically liberal fringe in the broader landscape of Chinese lawyers on the basis of empirical evidence.³⁰ Research on political values and ideology suggests that the Chinese population as a whole lean in a pro-regime, nationalistic direction.³¹ Chinese netizens appear to lean even more strongly in this direction.³² Pro-regime netizens also include Maoists and economic liberals; economic liberals share little in common

²⁴ See generally LIU & HALLIDAY, *supra* note 4.

²⁵ JONATHAN BENNEY, DEFENDING RIGHTS IN CONTEMPORARY CHINA 127–28 (2012).

²⁶ LIU & HALLIDAY, *supra* note 4, at 157.

²⁷ See generally Hong Lu et al., *Proceduralism, Political Embeddedness and Death Penalty Lawyers in China*, 238 CHINA Q. 353 (2019).

²⁸ Fu & Cullen, *supra* note 13, at 112; see generally Tin Muk Daisy Cheung, *Going Barefoot in the Middle Kingdom: A Preliminary Study of the Strategic Choices of Non-Licensed Weiquan Lawyers in Modern China*, 31 UCLA PAC. BASIN L.J. 1 (2013).

²⁹ See generally Li Kege, *Do Birds of a Feather Flock Together? Rights Protection (Weiquan) Lawyering in China*, 5 ASIAN J.L. & SOC'Y 1 (2018).

³⁰ But see, e.g., Michelson & Liu, *supra* note 1; John Wagner Givens, *Sleeping with Dragons? Politically Embedded Lawyers Suing the Chinese State*, 31 WIS. INT'L L.J. 734 (2014).

³¹ See generally Jennifer Pan & Yiqing Xu, *China's Ideological Spectrum*, 80 J. POL. 254 (2018); WENFANG TANG, POPULIST AUTHORITARIANISM: CHINESE POLITICAL CULTURE AND REGIME SUSTAINABILITY (2016).

³² See generally XU WU, CHINESE CYBER NATIONALISM: EVOLUTION, CHARACTERISTICS, AND IMPLICATIONS (2007); ONLINE CHINESE NATIONALISM AND CHINA'S BILATERAL RELATIONS (Simon Shen & Shaun Breslin eds., 2010); Angela Xiao Wu, *The Evolution of Regime Imaginaries on the Chinese Internet* J. POL. IDEOLOGIES (forthcoming); RONGBIN HAN, CONTESTING CYBERSPACE IN CHINA: ONLINE EXPRESSION AND AUTHORITARIAN RESILIENCE (2018); YA-WEN LEI, THE CONTENTIOUS PUBLIC SPHERE: LAW, MEDIA, AND AUTHORITARIAN RULE IN CHINA (2018).

ideologically with political or cultural liberals.³³ Perhaps we should expect that lawyers actively engaged in online discussions in 2004, at least to some extent, shared the political values of their latter-day nonlawyer counterparts. Or perhaps we should expect the opposite. After all, previous survey research on the relative importance Chinese people attach to economic and political rights suggests that political rights are, on average, far more important to lawyers than to nonlawyers, who privilege economic rights over political rights.³⁴

I will now turn to an analysis of the online conversations of thousands of lawyers across China in 2004 in the hopes that the ways in which they discussed topics of salience to them will shed light on their worldviews. In so doing, I will assess the extent to which the worldviews they expressed on the ACLA forum map onto their well-documented professional vulnerabilities stemming from their dealings with state actors—particularly in the criminal justice system—with clients, and with law firms.³⁵

II. DATA AND METHODS

The ACLA forum is an example of what is known as a “Web-based asynchronous threaded discussion forum,” which simply refers to an electronic space on a website where conversations unfold as linked chains or threads of messages separated by time as opposed to occurring in real time.³⁶ Discussions on such websites are also characterized by their hierarchical structure—layers of discussions, as some people reply to the original post while others reply to replies, to the replies of replies, etc. On threaded discussion forums, a discussion can split into separate discussions: a thread often divides into multiple threads, which is to say that conversations beget sub-conversations. Threaded discussion forums are contrasted with flat discussion forums, in which all replies are linked directly to the same original post.

After going live and growing gradually in the second half of 2002, the ACLA forum was never untethered from the state. Its content was

³³ See generally Ronggui Huang et al., *Beyond the Left-Right Spectrum: A Typological Analysis of Ideologues in China's Weibo Space*, 28 J. CONTEMP. CHINA 831 (2019); Angela Xiao Wu, *Ideological Polarization Over a China-as-Superpower Mind-set: An Exploratory Charting of Belief Systems Among Chinese Internet Users, 2008–2011*, 8 INT'L J. COMM. 2650 (2014).

³⁴ See generally Michelson & Liu, *supra* note 1.

³⁵ See generally Michelson, *supra* note 2; LIU & HALLIDAY, *supra* note 4.

³⁶ Laurie P. Dringus & Timothy Ellis, *Using Data Mining as a Strategy for Assessing Asynchronous Discussion Forums*, 45 COMPUTER EDUC. 141, 146 (2005).

monitored, its administrators occasionally removed politically sensitive content, and at times it was frozen in read-only mode, which prevented the publication of new content.³⁷ Before it went offline permanently in 2008, it could be accessed at <http://www.acla.org.cn/forum/> and <http://forum.acla.org.cn>. Both URLs have been preserved by the Internet Archive (https://web.archive.org/web/*/http://forum.acla.org.cn and https://web.archive.org/web/*/http://www.acla.org.cn/forum/). Beginning in the second half of 2005, the ACLA forum administrators implemented a policy requiring that every message be vetted and approved.³⁸

My archive includes 173,691 messages spanning October 3, 2002 through April 18, 2005. A “thread” refers to a discussion or set of discussions that are linked to a common source. An “original post” refers to the first message of a thread. An “original poster” refers to the user who posted the first message of a thread. I use “OP” to refer to both original posts and original posters. A “reply” refers to message posted in an existing thread established by an OP.

All 173,691 messages in my archive belong to 24,645 threads, of which 4,642 are isolates with no replies. That is, 149,046 messages in my archive ($173,691 - 24,645 = 149,046$) are replies to OPs. Of the sixty threads Liu and Halliday analyzed for their research (containing 834 messages), thirty-nine (containing 655 messages) were already in my archive.³⁹ Sida Liu generously shared the contents of their sample, which I incorporated into my archive.

The volume of non-redundant text in the subject lines and message bodies of all messages in my archive (i.e., after removing text in messages quoted from previous messages in the same thread) totals almost fifty million Chinese characters, Latin letters, and Arabic numerals. OPs account for about one-third of the total volume of text.

A. BULK DOWNLOADING

I downloaded messages from the ACLA forum using a web crawler. That is, I automated the process of bulk downloading messages from the ACLA forum by programming a web robot (Teleport Ultra), sometimes called a web spider, to save HTML pages according to rules I set. I dispatched the web robot twice; it collected two batches of messages.

³⁷ Halliday & Liu, *supra* note 5, at 77–78.

³⁸ *Id.*

³⁹ See generally Halliday & Liu, *supra* note 5; LIU & HALLIDAY, *supra* note 4.

In September 2004, the web robot downloaded 31,203 messages (in 2,273 threads). In May 2005, it collected 142,154 messages (in 22,351 threads). The remaining 334 messages belong to the twenty-one threads Sida Liu generously shared with me that were not originally in my archive. For the first batch, I instructed the web robot to follow links on HTML pages I created from search results of messages containing keywords of relevance to my research.⁴⁰ For the second batch, I instructed the web robot to begin its journey through the ACLA forum by following links to specific message boards, which generally (but not always) correspond to discussion topics. Every OP (and thus every thread) belongs to a message board. I sent the web robot to the 21 most popular message boards.⁴¹

In addition to downloading the contents of threads, the web robot also downloaded user profiles. Posting a message (either an OP or a reply) required being logged in with credentials from a user account. Very few messages in my archive (only 0.6%) are not linked to a specific user; these may be the messages of users whose accounts were “permanently deleted due to their radical posts.”⁴² User profiles were easy to collect because they were located at sequential URLs, obviating the need for the web robot to crawl through layers of hyperlinks.⁴³

⁴⁰ At the time, given my interest in professional regulation and unauthorized competition, I instructed the web robot to crawl through links starting from threads containing relevant key words, such as “unfair competition” (不正当竞争), “retired judges” (退休的法官), “fake lawyers” (假律师), “black lawyers” (黑律师, a euphemism for unlicensed lawyers), “peasant lawyers” (土律师), “steal clients” (抢当事人, 抢业务, and 抢活), etc. These links were mere starting points; the web robot followed and saved all links that satisfied URL-based inclusion and exclusion criteria. A complete list of keywords I used to establish starting points for the web robot’s first journey through the ACLA forum is available from me upon request.

⁴¹ Message boards I downloaded in the second batch include “legal advice” (法律咨询), “comments and suggestions” (建议意见), “marriage and family” (婚姻家庭), “company and securities” (公司证券), “labor disputes” (劳动争议), “contract law” (合同法), “rule of law commentary” (法治时评), “procedural justice” (程序正义), “miscellaneous” (灌水专区), “literary works” (原创文苑), etc. Because Sida Liu was a board manager, he had access to a password-protected “recycling bin” (回收站) board that contained politically sensitive content. He generously shared his password, which I gave to the web robot so it could download this board. A complete list of message boards I used to establish starting points for the web robot’s second journey through the ACLA forum is available from me upon request.

⁴² Halliday & Liu, *supra* note 5, at 77.

⁴³ For example, before the ACLA Forum was taken offline, the user profile of Sida Liu, whose username, or handle, was “红岩星,” was originally located at: <http://www.acla.org.cn/forum/showprofile.php?Cat=&User=27276> [<https://perma.cc/8K4V-U2YV>].

B. BUILDING THE ARCHIVE: A RELATIONAL DATABASE

The next step entailed using Perl scripts to extract or “scrape” relevant information from the HTML files the web robot successfully bulk downloaded. With respect to the contents of messages, I extracted the following information:

- unique message ID
- unique ID of the message to which this message is a reply
- message date
- message time
- discussion board
- message subject
- message body
- unique user ID
- user handle (username)
- user status
- user registration date
- total number of messages posted by user
- geographical location of user (as disclosed in the user’s profile)

With respect to the contents of user profiles, I extracted the following information:

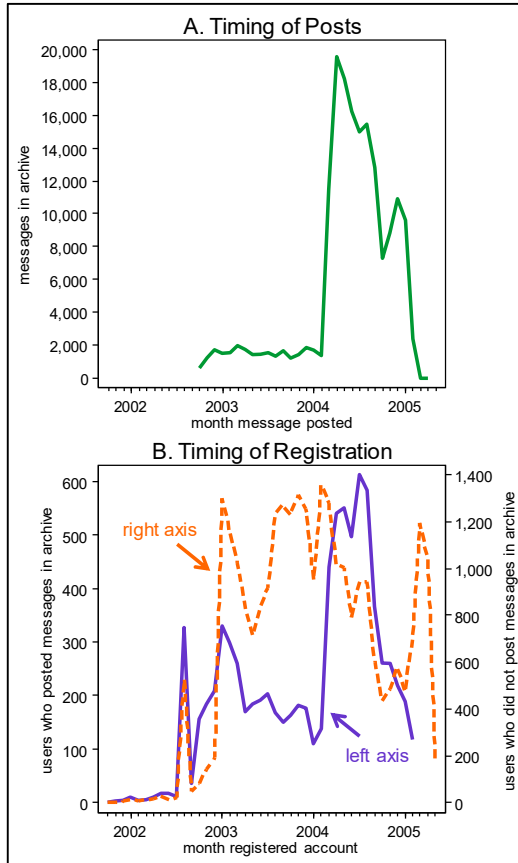
- user handle (username)
- user status
- user registration date
- total number of messages posted by user
- geographical location (open-ended)
- introduction (open-ended)
- occupation (open-ended)
- interests (open-ended)
- email address
- URL of website
- ICQ (an instant messaging and VoIP client)
- QQ (an instant messaging voice chat client)
- MSN (Microsoft Internet services user name; almost always a Hotmail email address)

We can see in Figure 1 that the contents of the both datasets are overwhelmingly from 2004.⁴⁴ Of all *messages* I downloaded, 80% were

⁴⁴ See *infra* Figure 1.

posted in 2004, and 2%, 11%, and 7% were posted in 2002, 2003, and 2005 respectively.⁴⁵ At the same time, the vast majority of *users* who posted the messages I downloaded registered their accounts in 2003 and 2004: 30% and 55% respectively.⁴⁶ Half of all users whose messages I did not download registered their accounts before 2004, while 37% and 13% registered their accounts in 2004 and 2005 respectively.⁴⁷

*Figure 1. Timing of Content Creation.*⁴⁸



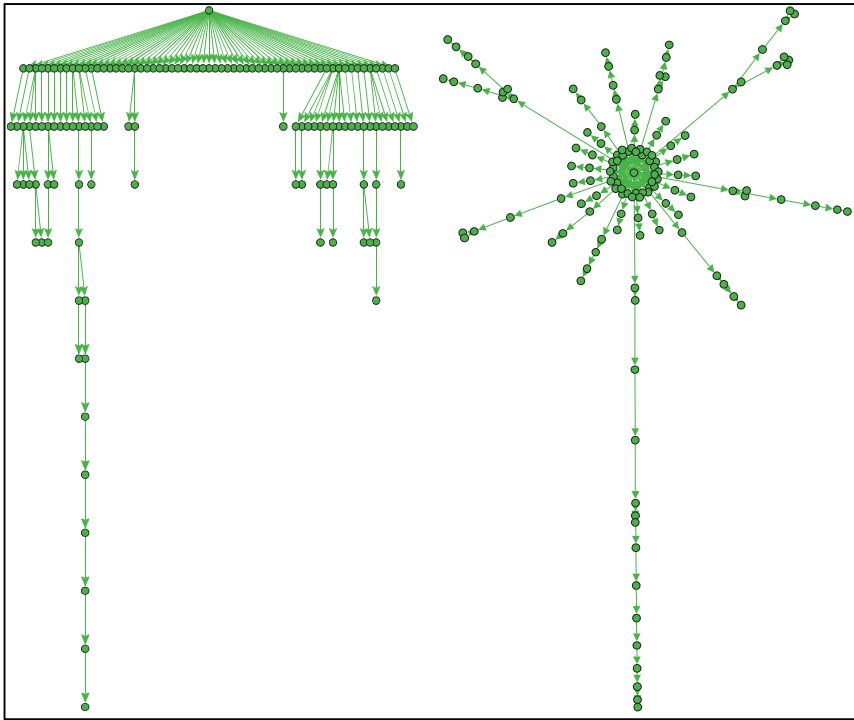
⁴⁵ *Id.*

⁴⁶ *Id.*

⁴⁷ *Id.*

⁴⁸ A zoomable color version of Figure 1 is available at <https://perma.cc/Q7NH-UPK3>. Figure A, *n*=173,691. Figure B, *n*=8,358 users who posted messages, *n*=28,163 users who did not post messages in the sample; total user profiles, *n*=36,521.

Figure 2. Two Visualizations of Thread #440419, “The Lawyer SOS Case.”⁴⁹



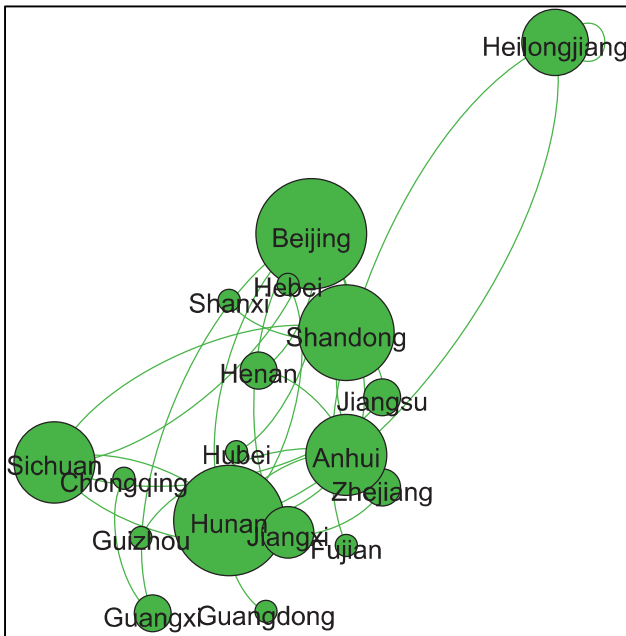
By merging the two datasets of messages and users, I created a multilevel relational database. Individual users can be easily linked according to discussion threads (common OPs), discussion boards, provinces, and cities. Because messages are linked to previous messages, entire discussion threads can be easily reconstructed. Figure 2 contains two methods of visualizing the same discussion thread (#440419),⁵⁰ dubbed “the lawyer SOS case” by Liu and Halliday because the OP was authored by a lawyer from Anhui Province being pursued by the police. The hierarchical structure of the discussion thread is immediately obvious as layers of replies in both layouts: many replies were made directly to the OP, but some replies spawned sub-discussions. In this regard, threaded discussions are genealogical. In addition to a common ancestor (a progenitor or forebearer from which all subsequent replies descend) shared

⁴⁹ A zoomable color version of Figure 2 is available at <https://perma.cc/RE2D-944B>.

⁵⁰ See *infra* Figure 2.

by all members of the thread, each message has a parent, which may or may not be the common ancestor. The genealogical analogy is imperfect, however, because each message has only one parent, users can reply to their own messages, and users can post multiple messages within a single thread. This particular thread contains 138 messages posted by 102 unique users: 137 replies to one OP.⁵¹ Of all 137 replies in this thread, sixty-one were to the OP and seventy-six were to messages other than the OP.⁵² The duration of time spanning the OP and the last message in the thread is 320 days.⁵³

*Figure 3. Geographical Ties in Thread #440419, “The Lawyer SOS Case.”*⁵⁴



Although not all users disclosed their geographical locations, we can nonetheless map geographical relationships using the limited

⁵¹ *Id.*

⁵² *Id.*

⁵³ *Id.*

⁵⁴ A zoomable color version of Figure 3 is available at <https://perma.cc/6N5D-HWZK>. Ties depicted in this figure are between messages and direct replies; it excludes indirect replies. Self-loops—ties between users in the same geographical location (including self-replies)—are permitted, and can be seen in the case of Heilongjiang (other self-loops are obscured by network nodes).

information they did disclose. The 102 unique users who participated in this same discussion thread about the “lawyer SOS case” disclosed twenty-one unique provinces. Figure 3 maps the geographical ties between users who posted messages and users who replied directly to their messages.⁵⁵ It shows the geographical diffusion of lawyers’ conversations in this thread.

Figure 4 depicts all geographical locations disclosed both by users who posted messages in the archive and by all users who registered accounts on the ACLA forum.⁵⁶ The sheer scale and scope of the community of participants fostered by the ACLA forum is palpable.

Limited occupational information disclosed by users suggests this community consisted largely of lawyers. Of all 8,365 registered users who posted messages in my archive, 1,341 provided an occupation, 67% of whom identified themselves as lawyers. The proportion of self-identified lawyers was higher among users who posted a larger share of messages. Among the most prolific users who disclosed an occupation, about 80% identified themselves as lawyers.

Figure 5 further boosts our confidence that lawyers accounted for the majority of users posting messages on the ACLA Forum.⁵⁷ The geographical distributions of both messages and the users who posted them correlate almost perfectly with the geographical distribution of the population of full-time lawyers at the time.⁵⁸ Conservatively estimating that two-thirds of registered users in my archive of messages were lawyers implies 5,500 unique lawyers, which is about 4% of all lawyers in 2004 (full-time, part-time, and interning). However, my archive represents only a small portion the entirety of the ACLA forum. The sum total of “total number of messages posted by user” in both the messages and the user profiles exceeds 577,000. An additional 28,160 users whose profiles I collected posted 22,864 messages that are not in my archive. My archive therefore consists of fewer than 30% of all 600,000 messages ever posted (not including those posted by an unknown number of users whose profiles were permanently deleted).

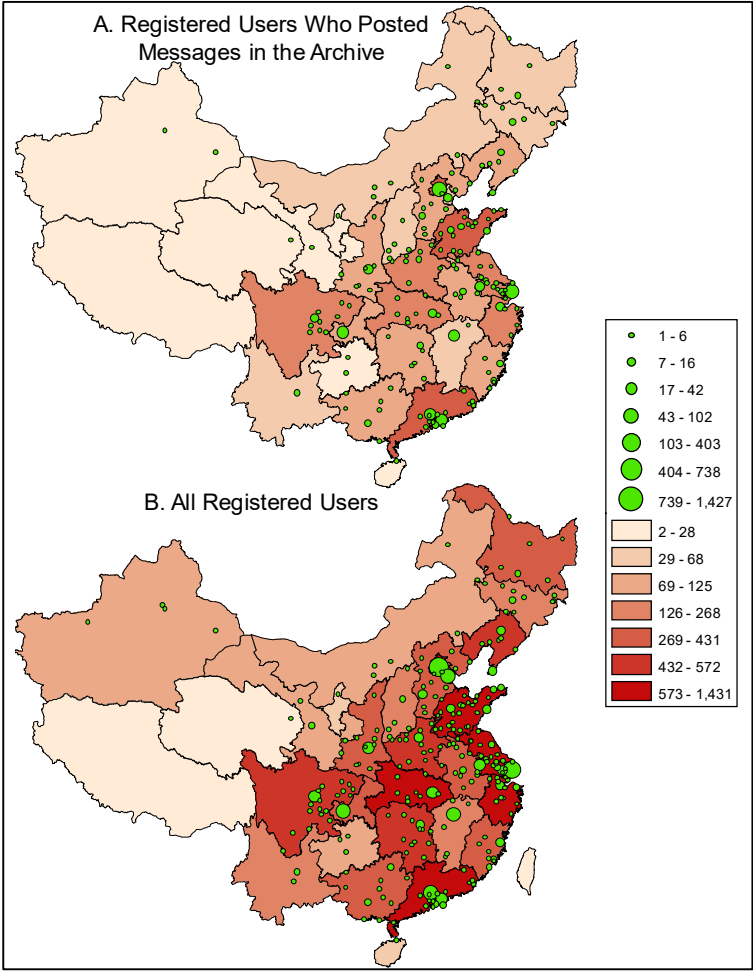
⁵⁵ See *supra* Figure 3.

⁵⁶ See *infra* Figure 4.

⁵⁷ See *infra* Figure 5.

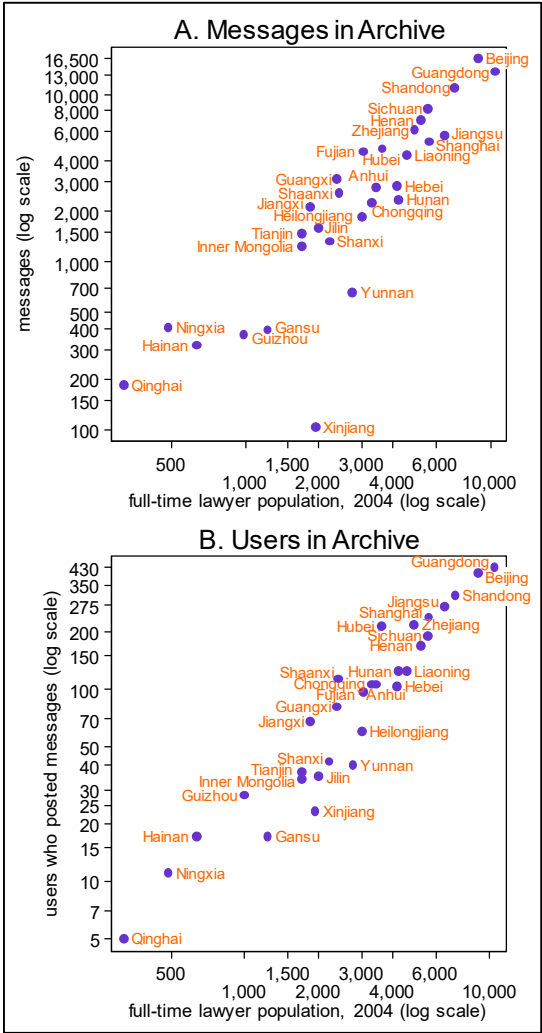
⁵⁸ *Chinese Yearbook of Lawyers 2004*, *supra* note 2, at 257, available at <https://perma.cc/8JAD-GF79>.

Figure 4. Geographical Locations Disclosed by ACLA Forum Users.⁵⁹



⁵⁹ A zoomable color version of Figure 4 is available at <https://perma.cc/B9AH-3X87>. Figure A contains 3,727 users who disclosed their provincial locations and 1,530 users who disclosed their municipal locations (163 unique cities). Figure B contains 12,803 users who disclosed their provincial locations and 4,102 users who disclosed their municipal locations (249 unique cities).

Figure 5. Distributions of Messages and Users in the ACLA Forum by the Distribution of the Total Population of Full-Time Lawyers.⁶⁰



Although the thread discussing the “lawyer SOS case” depicted in Figures 2 and 3 is one of the largest in the entire archive of 24,645

⁶⁰ A zoomable color version of Figure 5 is available at <https://perma.cc/9TWZ-A5N8>. Figure A, $n=114,119$ messages across 30 provinces, $r=.87$; Figure B, $n=3,727$ users across 30 provinces, $r=.95$. The official national population of full-time lawyers in 2004 was 106,325 (excluding Tibet, for which no data were reported).

threads,⁶¹ seventeen threads are even larger. The median number of messages per thread is 4 (mean=7.0, max=528). The median number of unique users contributing to threads is 3 (mean=4.8, max=218). The median number of disclosed provinces per thread is 2 (mean=2.7, max=28). The median duration of time over which discussions unfolded from beginning to end is 1 day (mean=9.1 days, max=631 days). Every message in the archive is unique; I deleted all duplicates.

C. TOPIC MODELLING

Having explained how I built the archive, I now describe how I made sense of its contents. The sheer volume of text in my archive of ACLA forum discussions prohibits manually reading and coding a sample much larger than Liu and Halliday's collection of sixty threads containing 834 messages.⁶² A sample of this size, however, accounts for only 0.2% of all threads and 0.5% of all messages in my archive. I therefore applied machine learning and natural language processing techniques to build a topic model. Topic modelling automates the process of identifying salient topics in unwieldy collections of documents such as my archive. In the terminology of topic modelling, my archive is the corpus, which consists of documents, which in turn consist of words. I faced a choice of analyzing individual messages as documents or aggregating messages into the threads of which they are part and treating each thread as a document. Because many messages are brief (sometimes only a few words), and in order to capture the flavor of the overall conversation, I chose the latter option, and treated the thread as the document for purposes of building a topic model. In this article I use "archive" and "corpus" interchangeably; I also use "thread" and "document" interchangeably. Nowhere in this article does "document" refer to "message." Messages form documents (discussion threads) but in this analysis are not themselves documents.

Before building a topic model, I first stripped out redundant text. I removed text from previous messages quoted in replies (both in subject lines and message bodies), which accounted for 14% of all text in the archive. The next step was to segment the text into words. Topic models are constructed from the words used to form documents. In most languages, including English, words are naturally identifiable by white space and punctuation. Because Chinese words are not segmented by

⁶¹ See *supra* Figure 2; see *supra* Figure 3.

⁶² See generally Halliday & Liu, *supra* note 5; LIU & HALLIDAY, *supra* note 4.

white space, I segmented all the text in my archive using the Stanford Word Segmenter (according to the Penn Chinese Treebank standard).⁶³

The topic model I present in this article was built without supervision. An unsupervised topic model is created without any prior information about the documents other than the words they contain. The model is “generative” in the sense that it is built iteratively through unsupervised machine learning; the machine gradually learns—and infers—which words are associated with unobserved, latent topics. The specific technique I applied is called Latent Dirichlet Allocation (“LDA”).⁶⁴

Not every word in the entire corpus of documents was a candidate for the model. A word was considered only if it met the following four criteria: (1) it appeared in at least two documents (threads), (2) it appeared in no more than 95% of documents (threads), (3) it was not filtered out as a meaningless word (by matching an item in a list of “stop-words”), and (4) it was among the top 600 most commonly used words remaining (i.e., it satisfied the first three criteria). Examples of “stop-words” excluded from the topic model are common subjects and objects (“I,” “you,” “he,” “she,” etc.), prepositions, articles, and so on. The list of stop-words I used contains 1,628 items, and incorporates established lists from Baidu, the Harbin Institute of Technology, and Sichuan University.⁶⁵

I constrained the analysis to a maximum of 20 topics for the sake of presentational parsimony—as we will see, even twenty topics are unwieldy! I also constrained the analysis to a maximum of 600 unique words in the so-called “bag-of-words matrix” on which the model is based.⁶⁶ Had I allowed the machine to identify the optimal number of topics (the point at which additional topics would not improve the predictive power of the model), my model would have included sixty-seven topics. Nonetheless, as we will see, a model of twenty topics successfully organizes discussions and their participants into clear clusters. Also for the sake of presentational parsimony, I limit my analysis

⁶³ Pi-Chuan Chang et al., *Stanford Word Segmenter v3.9.1* (2018).

⁶⁴ See generally David M. Blei et al., *Latent Dirichlet Allocation*, 3 J. MACHINE LEARNING RES. 993 (2003); Thomas L. Griffiths & Mark Steyvers, *Finding Scientific Topics*, 101 PROC. NAT’L ACAD. SCI. 5228 (2004); Fabian Pedregosa et al., *Scikit-learn: Machine Learning in Python*, 12 J. MACHINE LEARNING RES. 2825, (2011). I am grateful to Zuoyu Tian for writing the Python script that allowed me to use scikit-learn v0.19.1.

⁶⁵ I am grateful to Zuoyu Tian for helping to build the list of stop-words.

⁶⁶ Aneesha Bakharia, *Topic Modeling with Scikit Learn* (2016), <https://medium.com/mlreview/topic-modeling-with-scikit-learn-e80d33668730> (last visited Feb. 19, 2020), available at <https://perma.cc/LD3V-V2WW>.

to the 381 unique words among the top thirty words in each of the twenty topics in the model.

III. PLAN OF ANALYSIS

My plan of analysis consists of four steps. The LDA algorithm outputs two matrices that are the basis of all remaining analyses I present in this article. The first matrix is a *word-topic matrix*. It contains word weights for each of the twenty topics. Word weights measure the importance or salience of each word in a given topic, and “can be viewed as pseudocount [*sic*] that represents the number of times word *j* was assigned to topic *i*.”⁶⁷ We can use these weights to assess and visualize topics in at least two ways. First, we can use them to examine topics separately; they show us the mix of words that constitute and define topics. Thus, in the first step of my analysis, I will depict the distribution (according to word weight) of the top thirty words for each of the twenty topics in my model. Second, we can use the word weights to examine relationships between topics; they show us similarities/proximities and distances between topics. Thus, in the second step of my analysis, I will use word weights to depict the clustering of words and topics in what is known as a semantic network.

In order to construct a semantic network, I first transformed the 381×20 rectangular word-topic matrix containing word weights into a 381×381 square word-word matrix containing similarity coefficients. Because the resultant word-word matrix contains a lot of missing cells, I was unable to use common similarity measures. Of all 72,390 theoretically possible unique pairs of words $((381 \times 380)/2 = 72,390)$ in the model, 8,102 actually exist. Word weights are continuous variables that can be used to measure the strength of ties between words within and between topics in a semantic network. They can also be transformed into dichotomous variables to measure ties between words as either present or absent. In either case, popular similarity measures such as Euclidean distance (for continuous variables) and the Jaccard similarity coefficient (for binary variables) typically require dropping missing values.

My word-word similarity matrix therefore uses the Gower similarity coefficient for mixed data. Nodes (or vertices) in a semantic network are words, and edges (or ties) connecting the nodes are (modified)

⁶⁷ *Sklearn.Decomposition.LatentDirichletAllocation*, SCIKIT LEARN, available at <https://perma.cc/F657-C7VG>.

Gower similarity coefficients measuring the strength of the ties between words. Because edge weights can never equal zero, I slightly modified the Gower similarity coefficient.⁶⁸ I colored each node (word) in the semantic network according to its topic. I colored words that appear in multiple topics according to where they were most influential; the dominant topic of a given word is simply the topic with the maximum word weight in the word-topic matrix.

The second matrix is a *document-topic matrix*. It contains topic weights for each document (thread). Each document is distributed across multiple topics. Indeed, some documents are distributed across all twenty topics. Topic weights are transformed into proportions that sum to 1.0 within each document (with some rounding error). They are interpreted as the proportion of all words in the document from each topic. They thus allow us to identify a document's dominant topic, which almost always emerges conspicuously by a considerable margin; rarely if ever is there ambiguity about the topic to which a document belongs.

I will use the document's dominant topic for two purposes. First, I will assess the relative importance of each topic in the corpus of documents as a whole. Second, I will map a semantic network from the corpus of documents in order to assess the extent to which discussion threads and the individual users who contribute to them were segregated by topic. I mapped all networks using Gephi.⁶⁹ Because it is an undirected network (i.e., edges have no direction), I used the Yifan Hu layout to map the semantic network. This is the second step of the analysis.

In the third step of the analysis, I will depict the clustering of individual threads and the users who posted the OPs that spawned them. This exercise entails combining and connecting all applicable individual discussion threads. Whereas Figure 2 above contains one solitary

⁶⁸ My modification is the addition of 0.01 to the standard Gower similarity coefficient in order to prevent word-word ties of zero. The following is the equation I use to calculate similarities between words within a given topic:

$$S_{ij} = \left(1 - \frac{|x_i - x_j|}{\{max(x) - min(x)\}}\right) + 0.01,$$

where S_{ij} is the similarity between word i and word j . To illustrate, let us consider the degree of similarity between the words "lawyer" (律师) and "judge" (法官). They appear together only in Topic 2 (which I call "lawyers' plight"). Their word weights are 72,245 and 3,805 respectively. Meanwhile, the maximum and minimum weights among the top thirty words in this topic are 72,245 and 431 respectively. Thus, the similarity coefficient is $(1 - ((72,245 - 3,805) / (72,245 - 431))) + 0.01 = 0.057$. When a word pair appears in more than one topic, the similarity coefficient is simply the average of all its single-topic similarity coefficients. Values of this similarity measure range from 0.01 to 1.01.

⁶⁹ See generally Mathieu Bastian et al., *Gephi: An Open Source Software for Exploring and Manipulating Networks*, PROC. THIRD INT'L CONF. WEBLOGS SOC. MEDIA 361 (2009), available at <https://perma.cc/4PML-MYBQ>.

discussion thread, in this step of the analysis I form a network from almost 15,000 discussion threads.

I accomplished this task according to the following nine procedures. First, I excluded from the analysis all documents (threads) dominated by either of two topics of only marginal relevance to the legal profession (topics I call “general interest” and “forum-related”). I will return to this issue when I present my findings in the next section. Second, whereas I already determined the dominant topic of each *word* in the first step of my analysis (according to the word-topic matrix) and the dominant topic of each *document* in the second step (according to the document-topic matrix), in this third step I determined the dominant topic of each individual *user*. I calculated each user’s dominant topic simply as the modal (majority or plurality) topic among all the threads in which she or he appeared. In the event of ties, I randomly assigned a topic among the tied dominant topics. Third, I excluded from the analysis all OPs (regardless of topic) whose dominant topic is one of the two excluded from the analysis. Fourth, all discussion threads spawned by the same user emanate from a common OP node. In other words, I collapsed into a single node the OPs of every thread initiated by the same user. Fifth, I colored each OP according to the dominant topic of the user who posted it. Sixth, I colored each reply in each discussion thread according to the dominant topic of the thread. Because OPs are colored according to users and replies are colored according to document thread topics, we will see that OPs of one color generate threads of an assortment of colors. Even if an OP has a tendency to generate threads of a particular topic, she may occasionally initiate discussions on other topics. Edges connecting replies within threads are assigned weights of one. Seventh, I linked OPs according to the number of times their authors replied to each other in the corpus of discussion threads in the archive. That is, edges connecting OPs are assigned weights equal to the number of times one OP replied to the other. Edge weights connecting OPs are then multiplied by two to ensure they exert more influence on the layout of the network than replies within threads. Edge directions matter; a tie between two OPs is depicted with two edges if they both replied to each other at least once. If two OPs are connected by replies in only one direction, the tie is depicted with a single edge. Eighth, in order to facilitate the visual identification of influential users (prolific posters), I scaled OP nodes according to the share of all messages in the corpus they posted. Ninth, because it is a directed network (edges have directions), I used the ForceAtlas2 layout to map the network of discussion threads.

In the fourth and final step of the analysis, I will illustrate the topics qualitatively with examples of discussion threads from selected topics. I selected examples according to size—that is, I chose to illustrate selected topics with particularly large discussion threads as well as discussion threads initiated by particularly prolific OPs. These are the threads that exerted the most influence on specific topics and the ACLA forum more generally. I also selected examples randomly to ensure I drew from a representative cross-section of documents.

IV. FINDINGS

I present my findings in four steps according to the plan of analysis I described in the previous section. The first step is to display the twenty topics and the mix of words that define them. The second step maps relationships between words and topics as a semantic network. The third step maps the majority of discussion threads in the corpus. The fourth and final step fleshes out selected topics with examples of discussion threads.

A. STEP ONE

Table 1 contains a list of all twenty topics in the model with their corresponding labels and colors used to depict them in the remainder of this article.⁷⁰ It serves as a color reference for the subsequent figures. All tables and figures are available online, in color where applicable; see each table's corresponding footnote for permalinks. It also contains the share of all messages in the corpus. The basis of my labels will become apparent when we see the words that define each topic (later in this step) and examples of discussion threads that illustrate selected topics (in the fourth step). For now, the key finding is the overwhelming importance of the "general interest" topic in the archive as a whole. This topic alone accounts for one-quarter of all messages in the archive.⁷¹

⁷⁰ See *infra* Table 1.

⁷¹ *Id.*

*Table 1. Topic Distribution and Color Attributes*⁷²

Topic	Color	Color Description	Share of Archive
1. general interest		purple	24.6%
2. lawyers' plight		apricot	7.3%
3. political liberalism		navy	7.2%
4. contracts		teal	6.4%
5. nationalistic		brown	5.8%
6. forum-related		yellow	5.7%
7. criminal justice		green	5.6%
8. marriage and divorce		black	5.2%
9. traffic torts		olive	4.4%
10. commerce/investment		lime	4.1%
11. litigation		lavender	3.5%
12. education		cyan	3.5%
13. labor		maroon	3.3%
14. court decisions		blue	2.8%
15. professional regulation		pink	2.3%
16. health and medicine		orange	2.0%
17. law firms		white	1.9%
18. administrative		red	1.8%
19. government		mint	1.5%
20. intellectual property		grey	1.2%
Total			100.1%
Archive (messages)			173,691

Comparing topics inferred by the model with message board topics helps validate the model. Substantive legal topics inferred by the model correspond to board topics on specific areas of law. For example, 57% of all threads in Topic 13 (labor) were posted to the “labor conflict” board.⁷³ Likewise, 45% of all threads in Topic 12 (education) were posted

⁷² A color version of Table 1 is available at <https://perma.cc/ND9Q-L7J2>.

⁷³ See *infra* Table 1; *supra* note 41.

to the “judicial examination” board.⁷⁴ Similarly, 31% of all threads in Topic 7 (criminal justice) were posted to the “criminal law” board.⁷⁵ At the same time, 32% of all threads in Topic 10 (commerce/investment) were posted to the “company and securities” board.⁷⁶ Finally, 25% of all threads in Topic 8 (marriage and divorce) were posted to the “marriage and family” board.⁷⁷

Message board names, however, are no substitute for the topics inferred by the model. On the contrary, they obscure important distinctions between topics. For example, the “legal advice” board contains a hodgepodge of threads that belong to a variety of inferred topics, including not only the ones mentioned above (32% of threads on “labor conflict,” 14% of threads on “education,” 28% of threads on “criminal justice,” 39% of threads on “commerce/investment,” and 37% of threads on “marriage and divorce”), but also a large share of threads classified by the model as “administrative,” “court decisions,” “health and medicine,” “contracts,” “litigation,” “government,” “traffic torts,” and “intellectual property.” Meanwhile, threads classified by the model as belonging to Topic 3 (political liberalism) and Topic 5 (nationalistic) are both concentrated in the “rule of law commentary” board (55% and 68% of threads in each respective topic).⁷⁸ Finally, “general interest,” by far the largest topic in the corpus, is spread fairly evenly across a number of message boards, including “literary works,” “legal advice,” “rule of law commentary,” and “miscellaneous small talk.”⁷⁹ In short, compared to the message board topics, the topics inferred by the model are more discrete and analytically helpful.

Figure 6 contains word clouds for each topic inferred by the model.⁸⁰ For reference purposes, the Appendix contains every word in every topic listed in order of importance within the topic, their corresponding original Chinese words, and additional topics containing the same words. The words in Figure 6 are scaled according to the word weights in the word-topic matrix. Topic colors correspond to those in Table 1.

⁷⁴ See *supra* Table 1; *supra* note 41.

⁷⁵ See *supra* Table 1; *supra* note 41.

⁷⁶ See *supra* Table 1; *supra* note 41.

⁷⁷ See *supra* Table 1; *supra* note 41.

⁷⁸ See *supra* Table 1; *supra* note 41.

⁷⁹ See *supra* Table 1; *supra* note 41.

⁸⁰ See *infra* Figure 6.

Figure 6. Word Clouds for Inferred Topics.⁸¹



⁸¹ A zoomable color version of Figure 6 is available at <https://perma.cc/785G-V7TT>. Sum total of all topics equals 100.1% owing to rounding error. $n=24,645$ documents (discussion threads) containing $n=173,691$ messages. Words are scaled according to word weights.

Word clouds in Figure 6 depict the top thirty words in each topic scaled according to their importance or salience. The 600 words in total (30 words $\times$ 20 topics) consist of 381 *unique* words. Many words appear in more than one topic. The word that appears in the most topics is “problem” (问题, in fourteen topics), followed by “stipulation” (规定, in nine topics), “law” (法律, in eight topics), “ought” (应当, seven topics), “think” (认为, seven topics), “whether” (是否, six topics), “nation” (国家, six topics), “work” (工作, six topics), “time” (时间, six topics), and “people” (人民, six topics). A word’s importance, however, is not determined by the number of topics in which it appears. Because they appear in so many topics, these words figure prominently in no topics. The most influential words, by contrast, are “lawyer” (律师, Topics 2 and 6), “company” (公司, Topic 10), “court” (法院, Topics 11 and 14), “contract” (合同, Topics 4 and 13), and “China” (中国, Topics 5 and 6).

The lexicon of a topic provides clues about the nature of the discussions they contain. Of particular importance to my motivating research questions is Topic 3, which contains politically liberal discourse reflected in the use of words such as “rule of law,” “democracy,” “right,” “reform,” and “constitution.” Topic 3 is evidence of “a liberal moment,”⁸² albeit a limited one. To characterize the ACLA forum as a whole as colored predominantly by politically liberal discourse would be misleading. The corpus of discussion threads is not awash with political liberalism. Political liberalism was clearly present and visible, but was far from ubiquitous. Politically liberal discourse was confined to a subset of discussion threads and users. Lawyers did not always frame their bitter complaints about various sources of professional vulnerability (Topic 2)—including unauthorized competition (Topic 15 on professional regulation)—in the language of political liberalism.

A contrasting but similarly important type of political discourse is visible in Topic 5, which contains nationalistic discourse focused on foreign meddling, particularly on the part of the US, historical and contemporary conflicts with Japan, and the status of Taiwan. Both types of political discourse coexist with more mundane, practitioner-oriented discussions about contracts (Topic 4); criminal justice (Topic 7); marriage and divorce (Topic 8); traffic torts (Topic 9); commerce, investment, finance, and real estate (Topic 10); and labor conflict (Topic 13). To be sure, some of these areas of practice, most notably criminal defense and

⁸² See generally Halliday & Liu, *supra* note 5.

labor, were animated by challenges that lawyers understood and discussed in political terms. Most of the time, however, when users of the ACLA forum discussed specific areas of legal practice, they did so without invoking political liberalism. In terms of share of the corpus, practice-oriented topics were far more prevalent and important than the two political topics combined.

Meanwhile, political and practitioner-oriented topics alike were overwhelmed by general interest topics unrelated to law (Topic 1), such as literature. Similarly, forum-related discussions (Topic 6) concern matters, which are sometimes technical, related to the administration of the ACLA forum itself. Taken together, general interest and forum-related discussions account for almost one-third of all content in the entire archive of discussion threads.

B. STEP TWO

Figure 7 is a semantic network that shows the strength of ties between words. By coloring nodes according to the topic in which they have the most influence (the largest word weight for each word determines its dominant topic), we can also assess similarities between topics. As we already know, the 600 words in Figure 6 (30 words per topic $\times$ 20 topics = 600 words) are reduced to 381 unique words in the semantic network depicted in Figure 7 because many of the top thirty words in each topic span multiple topics. (The 600 words in Figure 6, many of which are redundant, and from which the 381 unique words in Figure 7 originate, should not be confused with the 600 unique words used to create the model.)

Topic 1 (general interest) retains more words than any other topic. Among its top thirty words, only “problem” (问题) is defined as belonging to another topic. The word “problem” appears in Topic 3 (political liberalism) more than in any other topic. The dominant topic of each word is determined according to the maximum word weight in the word-topic matrix, which in the case of the word “problem” belongs to Topic 3 (political liberalism). Topic 3 (political liberalism) has the second largest collection of words in Figure 7. It retained twenty-five of its top thirty words, losing only “law” (法律, to Topic 15, professional regulation), “China” (中国, to Topic 5, nationalistic), “right” (权利, to Topic 8, marriage and divorce), “people” (人民, to Topic 11, litigation), and “think” (认为, to marriage and divorce). Topic 17 (law firms) retained twenty-three of its top thirty words. On the other end of the spectrum are

Topic 13 (labor, which includes twelve words) and Topic 2 (lawyers' plight, which includes thirteen words). The thirteen words, in order of importance (from most to least), that belong predominantly to Topic 2 (lawyers' plight), reflect how lawyers talk about their professional vulnerabilities: "lawyer" (律师), "firm" (事务所), "practice" (执业), "fee" (收费), "profession" (职业) "association" (协会), "industry" (行业), "lawsuit" (官司), "maintain" (维护),⁸³ "rights and interests" (权益), "stage" (阶段), and "environment" (环境).

It is only natural that words tend to cluster within topics. After all, edge weights (the strength of ties between words) are calculated primarily according to the joint presence of words within topics. Of greater significance than the locations of individual words are the locations of the topics they form. Topics in the semantic network cluster in an orderly fashion along several dimensions. First, the topics that have the least to do with law and legal practice (Topic 1 in purple, general interest, and Topic 6 in yellow, forum-related) are on the periphery of the network in the lower-left quadrant. Second, topics related to law firms (Topic 17 in white), professional regulation (Topic 15 in pink), and education (Topic 12 in cyan) cluster together in the upper-left quadrant. Topics 1, 6, and 17 (general interest, forum-related, law firms, respectively) are the most isolated of all topics. Their vocabularies overlap the least with other topics. Among all 381 words in all twenty topics, words appeared on average in 2.7 topics. Words in Topic 17 (law firms), by contrast, appeared on average in only 1.7 topics. Just behind this topic is Topic 1 (general interest), in which words appeared on average in only 1.8 topics. In the third most isolated topic, Topic 6 (forum-related), words appeared on average in 2.3 topics.

⁸³ "Maintain" is one of many correct translations of this word but is not terribly suitable in this context. More appropriate English words—such as "protect" and "defend"—were already taken by "保护" (Topic 8, marriage and divorce) and "辩护" (Topic 7, criminal justice) respectively.

Figure 7. A Semantic Network of the ACLA Forum.⁸⁴

Meanwhile, the two political topics, political liberalism (Topic 3 in navy) and nationalistic (Topic 5 in brown), cluster closely together. Adjacent to these political topics are lawyers' plight (Topic 2 in apricot) and labor conflict (Topic 13 in maroon). Occupying the right side of the semantic network and surrounding the political topics are a wide array of practitioner-oriented topics, including Topic 4 (in teal, contracts), Topic 10 (in lime, commerce/investment), Topic 7 (in green, criminal justice), Topic 8 (in black, marriage and divorce), Topic 9 (in olive, traffic torts), Topic 11 (in lavender, litigation), and Topic 14 (in blue, court decisions). The least prevalent topics of government (Topic 19 in mint) and intellectual property (Topic 20 in grey) cluster together towards the bottom of the semantic network.

In a general way, topics are divided between politics and professional difficulties on the left and practice-oriented matters on the right. Despite their dramatic ideological differences, politically liberal and nationalistic topics share a lot in common lexicologically. Indeed, they share ten words: "nation" (国家), "China" (中国), "government" (政府), "problem" (问题), "politics" (政治), "economy" (经济), "people" (人民), "development" (发展), "think" (认为), and "become" (成为). In terms of their lexicons, these two topics share far more in common with each other than they do with practice-oriented topics.

In this step we have seen important lexicological similarities and differences between topics. Reducing the entire corpus of discussion threads to their 381 most important words and mapping their semantic relationships reveals how discussion topics cluster according to political and practice-oriented content. Despite their sharp ideological differences, politically conservative and politically liberal discourses, both of which were prevalent on the ACLA forum, shared a common (if not identical) vernacular. In the next step we will see if these relationships between topics persist at a more granular level.

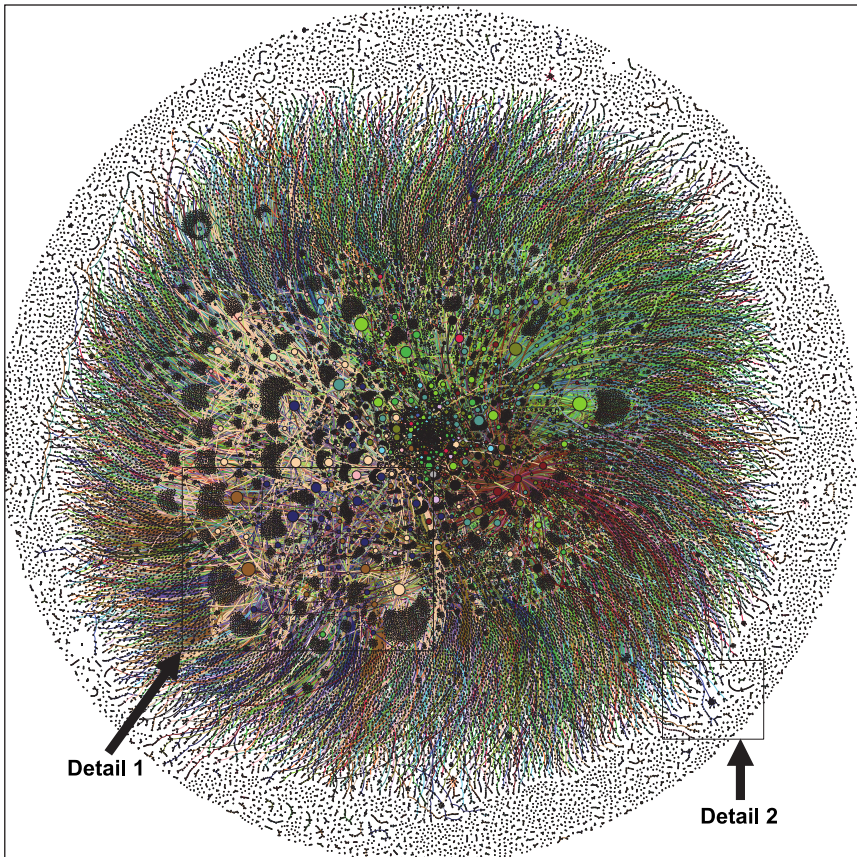
C. STEP THREE

In the previous step I mapped relationships between the 381 unique words that constitute topics in the model. In this step I attempt to map relationships between topics while retaining the full detail of the corpus. I will now map relationships between topics at a more granular level by linking individual messages, the discussion threads to which they belong, and the OPs that spawned them. Whereas Figure 2 depicts the structure of a single discussion thread (colored in green because it belongs

to Topic 7, (criminal justice), Figure 8 depicts and links a larger corpus of discussion threads.

In order to reduce noise, I excluded two large and distracting topics: Topic 1 (general interest) and Topic 6 (forum-related). I dropped not only all messages belonging to discussion threads in these two topics, but also the OPs of users defined as belonging predominantly to these topics. Messages that were orphaned as a consequence of the above exclusion rules were also dropped unless they were connected to at least one reply—that is, isolates are excluded. The remaining corpus depicted in Figure 8 consists of 103,543 messages posted by 6,865 unique users in 14,630 discussion threads. Among these 14,630 discussion threads, 12,385 are intact and 2,245 were decapitated (and are therefore partial) by excluding OPs who belonged predominantly to the two excluded topics. (The messages of users who belonged predominantly to the two excluded topics were excluded only if they were OPs; their replies in threads that belonged to non-excluded topics were retained in the analysis.) As I described in the previous section, I consolidated into single nodes all OPs of individual users. For this reason, the remaining corpus of 103,543 messages is reduced to 94,883 nodes. That is, all threads generated by an individual user emanate from a single node. As I also described in the previous section, OPs are linked according to the number of direct replies they made to and received from other OPs in the remaining corpus. The 94,883 nodes in Figure 8 are linked by 136,115 ties (edges). Also recall that OP nodes are scaled according to the share of the entire corpus the OP contributed.

In the online version of Figure 8 (available in color at <https://perma.cc/T5AT-B3DY>), it is possible to zoom in on sections of the network, and even to trace the threads of specific OPs. For the convenience of print readers, I also provide two inset maps—Figure 9 (Detail 1) and Figure 10 (Detail 2). The detailed portions of Figure 8 depicted in Figures 9 and 10 also contain the user IDs and usernames of OPs. Figure 9 (Detail 1) shows a section of the network dominated by a high density of discussion threads on political topics, Topic 3 (political liberalism) and Topic 5 (nationalistic). Figure 10 (Detail 2) shows a section on the periphery of the network containing the tails of long threads, decapitated threads, and threads spawned by OPs with no connections to any other OPs.

Figure 8. Discussion Threads Linked by OPs.⁸⁵

In network visualizations, nodes are situated according to forces of attraction and repulsion that minimize the energy of the system as a whole. The most prolific and influential OPs are depicted as larger nodes in the network. The threads they spawned appear as surrounding swarms. Waves of swarms extending outwards from OP nodes reflect the hierarchical nature of discussion threads (discussed earlier and illustrated with Figure 2). Longer discussion threads extend outwards like stalks of

⁸⁵ A zoomable color version of Figure 8 is available at <https://perma.cc/T5AT-B3DY>. It depicts 14,630 discussion threads containing 103,543 messages posted by 6,865 unique users. They are depicted as 94,883 nodes (136,115 edges) because all threads initiated by a single OP stem from the same node (multiple OP nodes were consolidated into single OP nodes). Colors (available in the online version) correspond to topics in Table 1. ForceAtlas2 layout.

wheat. At the core of the network are less frequent OPs who spawned fewer discussion threads; OPs of smaller mass were less able to resist the force of gravity pushing them towards the center of the network. At the outer perimeter of the network are threads whose OPs are unconnected to any other OPs⁸⁶ and decapitated threads whose OPs were excluded because they belonged predominantly to the excluded “general interest” and “forum-related” topics. Because they are not linked to any other threads, the locations of these disconnected threads in the network are of no substantive significance.

Note the remarkable consistency between Figures 7 and 8 in the clustering of political topics and practice-oriented topics. “Nationalistic” (Topic 5, brown) and “political liberalism” (Topic 3, navy) cluster in the lower-left half of Figure 8, whereas the various practice-oriented topics cluster in the upper-right half of Figure 8. Also no different from Figure 7, “lawyers’ plight” (Topic 2, apricot) and “professional regulation” (Topic 15, pink) are aligned more closely with the two political topics than with the practice-oriented topics. Whereas Figure 7 shows that the *words* used in both political topics overlapped, Figure 8 shows that the *people* who contributed their words to both political topics also overlapped. The users who participated in “nationalistic” discussions were often the same users who participated in “political liberalism” discussions, and they used similar vocabularies when they contributed to both topics. Recall that each OP is colored according to the predominant topic of the user who posted it, whereas replies are colored according to topic of the discussion thread.

Upon closely inspecting Figure 9, one can see that any given OP initiated discussion threads on a variety of topics. OPs classified as “nationalistic” (such as user 44225:日月盈) also spawned discussion threads classified as “political liberalism.” The opposite is also true. Because he contributed predominantly to “political liberalism” topics, Sida Liu’s OP node was colored navy (27276:红岩星). At the same time, he spawned eight discussion threads classified by the topic model as “nationalistic.” Perhaps users shifted seamlessly between nationalistic and politically liberal discourses. In their minds there may have been no contradiction between the two. Another possibility is that adherents of both persuasions debated one another. A third possibility is that the topic model is an overly blunt tool. Although my reading of Liu’s eight discussion threads classified as “nationalistic” lends support to the third

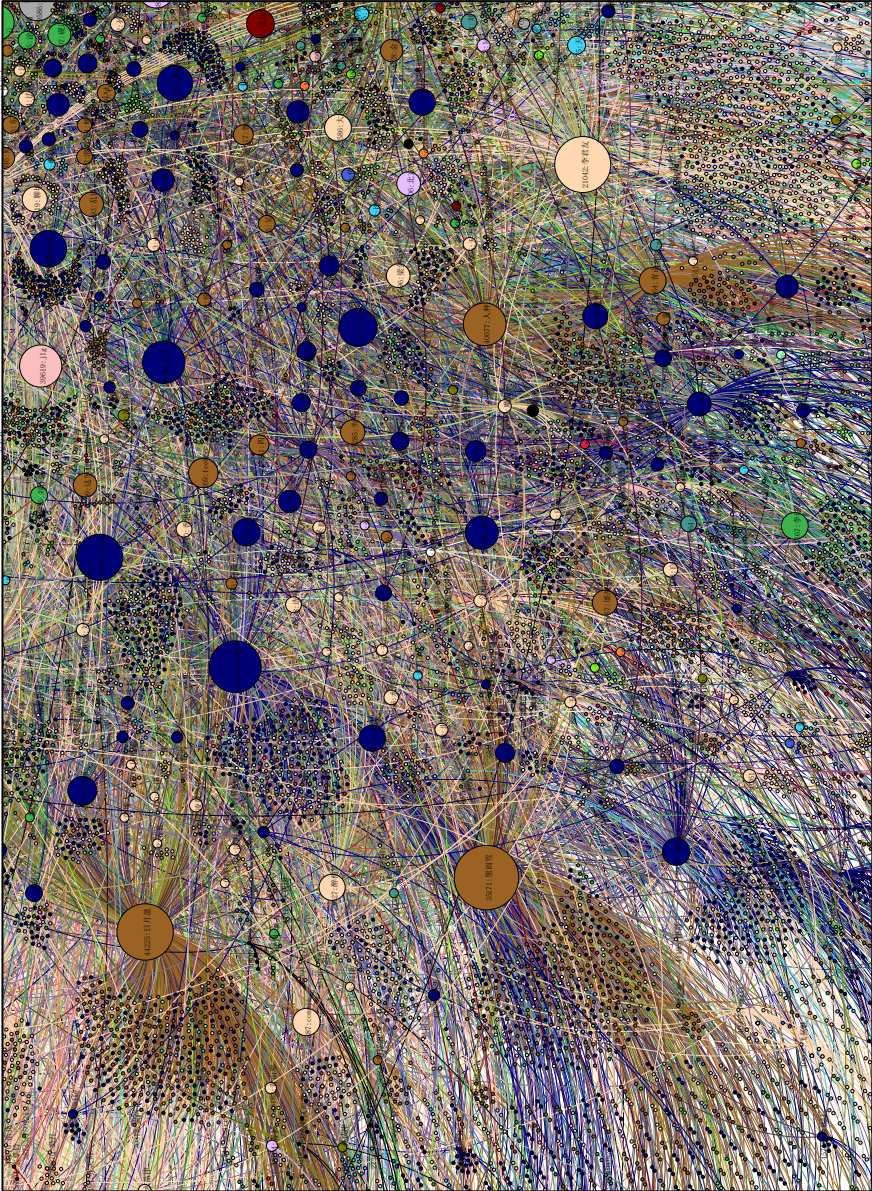
⁸⁶ This happens when two conditions are satisfied: (1) the only threads in which an OP appears are those spawned by the OP and (2) none of the users who replied to the OP in any thread spawned by the OP was an OP of another thread.

possibility, namely that distinctions in the topic model between “political liberalism” and “nationalistic” can be imprecise, more thoroughly adjudicating between these three possibilities must be left to future research.

Figure 8 as a whole and Detail 1, depicted in Figure 9, also bring into high relief the influential role of particularly prolific users, whom we can think of as super-users. By far the most productive contributor to the corpus was user 43259:听雨轩. This user contributed predominantly to the “general interest” topic and is therefore excluded from Figure 8 as an OP. However, this user’s replies—all 543 of them—are depicted in Figure 8. Overall, this user contributed 859 OPs and 2,203 replies to the archive as a whole in 188 days (spanning a period of 326 days between his first and last messages). On the days this user contributed posts, his or her average daily contribution was 16.3 posts, each averaging about 680 Chinese characters in length (meaning he or she posted an average of over 10,000 Chinese characters every day he or she posted anything!). This user is an extreme outlier. The extreme length of this user’s messages is a function of the nature of the “general interest” topic to which he or she predominantly contributed. Users across topics often shared and discussed media reports on salient topics. OPs classified as “general interest” often contained short stories, poems, and other literary works.

In Figure 8, the forty most prolific posters—who accounted for only 1% of all 6,865 users in Figure 8—contributed a whopping 19% of all OPs and 18% of all replies. Half of all discussion threads in Figure 8 were spawned by the top 8% of all OPs. One-fifth of all discussion threads in Figure 8 were spawned by the top 1% of all OPs. Replies to OPs were even more concentrated: half of replies in Figure 8 were posted by the top 4% of all repliers. Such a high degree of concentration is not unique to China; it is an intrinsic feature of all online discussions everywhere. Research suggests that nowhere in the world do online discussion forums facilitate widely participatory deliberative practices.⁸⁷

⁸⁷ See generally Davy Janssen & Raphaël Kies, *Online Forums and Deliberative Democracy*, 40 ACTA POLITICA 317 (2005); Scott Wright, *The Impact of “Super-Participants” on Everyday Political Talk*, 17 J. LANGUAGE POL. 155 (2018); Steven M. Schneider, *Expanding the Public Sphere through Computer-Mediated Communication: Political Discussion about Abortion in a Usenet Newsgroup* (May, 2 1997) (Ph.D. dissertation, Massachusetts Institute of Technology); RAPHAËL KIES, *PROMISES AND LIMITS OF WEB-DELIBERATION* 43 (2010).

Figure 9. Close-Up of Detail 1 in Figure 8 (Rotated 90°).⁸⁸

⁸⁸ A zoomable color version of Figure 9 is available at <https://perma.cc/A3QT-J8QN>. This zoomed-in view of Detail 1 in Figure 8 also includes node labels showing user IDs and user handles (usernames) of OPs.

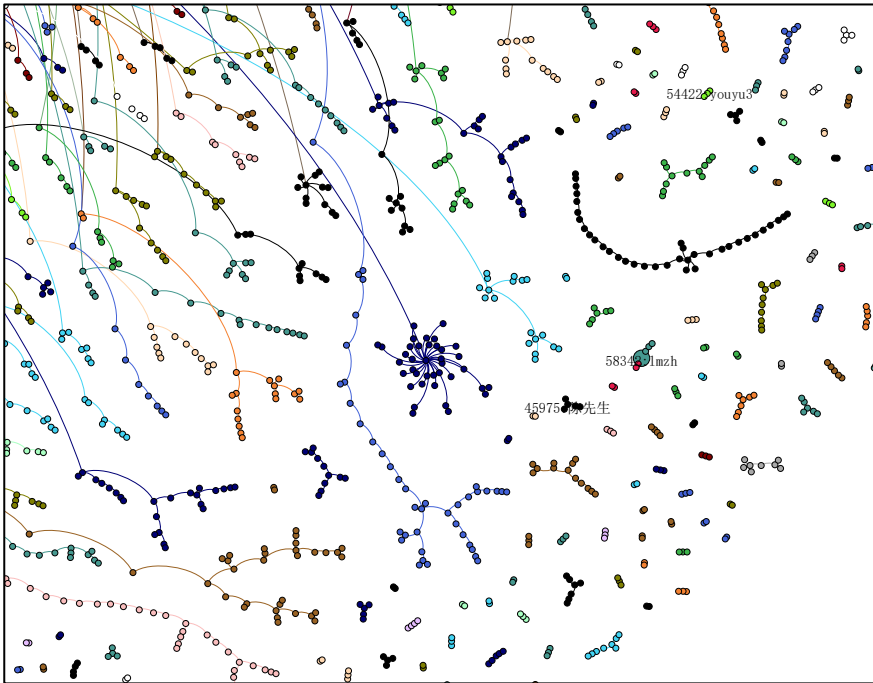
Figure 10. Close-Up of Detail 2 in Figure 8.⁸⁹

Figure 11A depicts the concentration curves for all users who appear in Figure 8.⁹⁰ Individual users at the very top of the concentration curves can be identified in Figure 11B. Only the top forty contributors are labeled in Figure 11B (i.e., the forty most prolific posters discussed above). Figure 11 therefore serves the purpose of revealing the source of prominent and influential nodes in Figure 8. For example, Figure 11B shows that the most prolific OP in Figure 8 was 沉舟侧, who is depicted as a lime circle because he contributed predominantly to the “commerce/investment” topic.⁹¹ This particular user appears in Figure 8 as the lime node to the right of the center (the largest node in the network). This one user initiated 215 discussion threads (OPs) and made 285 replies,

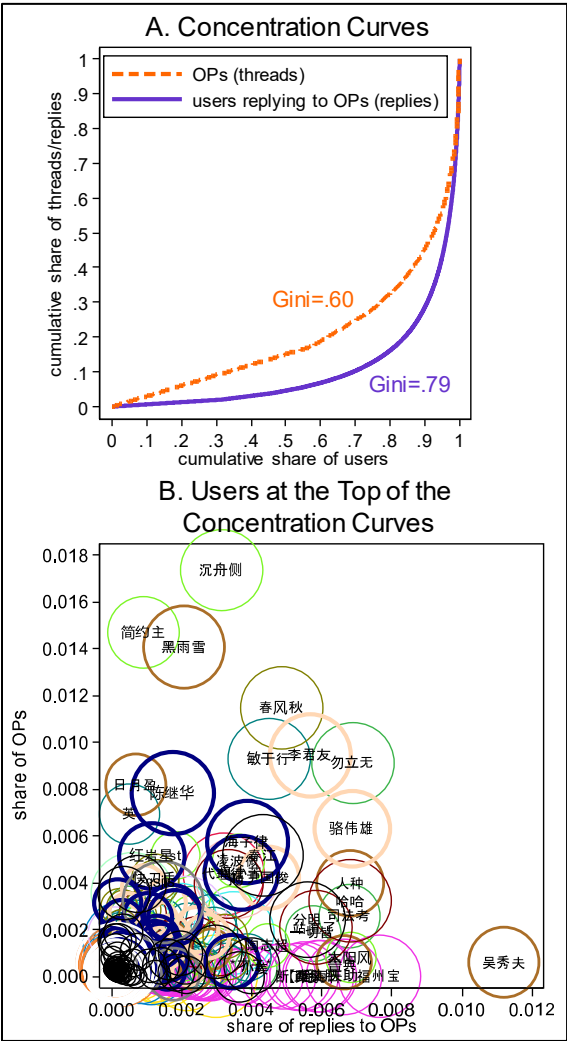
⁸⁹ A zoomable color version of Figure 10 is available at <https://perma.cc/87K3-QUM3>. This zoomed-in view of Detail 2 in Figure 8 also includes node labels showing user IDs and user handles (usernames) of OPs.

⁹⁰ See *infra* Figure 11A. Concentration curves do not vary wildly by discussion topic.

⁹¹ See *infra* Figure 11B. For purposes of the analysis and presentation of findings I truncated all user handles to three characters. This particular user’s full handle was “沉舟侧畔.”

all of which are depicted in Figure 8. Considering only the messages contained in Figure 8, this user, over a period of 334 days spanning the first and last messages, posted an average of 3.4 messages on each of the 145 days the user posted anything. This user’s posts tended to be short, averaging a Tweet-sized 150 Chinese characters.

Figure 11. Concentrations of OPs and Replies among Users.⁹²



⁹² A zoomable color version of Figure 11 is available at <https://perma.cc/FY8S-SLUG>. There are 6,860 unique users in both figures: 3,725 OPs and 5,783 repliers. Since many OPs are also repliers,

In Figure 11B, the two most productive “nationalistic” OPs (in brown circles) are 简约主 and 日月盈. Both users appear in Figure 8 as the largest brown nodes inside Detail 1. They are therefore even more visible in Figure 9 (which is a close-up view of Detail 1). 简约主 initiated 174 discussion threads and 189 replies in Figure 8; 日月盈 contributed 101 discussion threads and sixty-four replies in Figure 8. As my final illustration, Sida Liu (whose username was “红岩星”) has the distinction of being the fifteenth largest contributor of OPs in Figure 8 (out of all 3,725 users who contributed OPs in the figure). As a user who contributed predominantly to the “political liberalism” topic, Liu launched sixty-four discussion threads and contributed 105 replies depicted in Figure 8. He is easily identifiable as the navy node at the very top of Detail 1 in Figure 8 and therefore also at the very top of Figure 9.

Let us take stock of the empirical findings presented thus far. First, participants in online discussions on the ACLA forum were very likely to talk about matters unrelated to legal practice. Second, participants were more likely to discuss practice-oriented topics than to discuss political topics. Third, when they did discuss political topics, they were about as likely to discuss nationalistic topics as they were to discuss political liberalism. Fourth, users who discussed political liberalism also discussed nationalistic topics; users spanned both categories and interacted with one another. Fifth, and finally, online discussions on the ACLA forum across all topics, including “political liberalism,” were overwhelmingly concentrated among a small subset of users.

D. STEP FOUR

We now have a general picture of the topics ACLA forum participants—most of whom were lawyers—discussed in 2004, their distribution, and how they related to each other in terms of the discussion participants and the lexicons they used. Let us now zoom in to specific conversations in an effort to get a better grip on the precise nature of the discussion content. As I mentioned in the previous section, I selected examples both purposively (large threads, particularly those contributed by super-user OPs with large and influential nodes in Figure 8) and

some users are in both lines in Figure A (i.e., the sum of the users in both lines exceeds the total number of unique users because users who are both OPs and repliers are double-counted). Figure A depicts Lorenz curves. Colors in Figure B correspond to topics in Table 1. Marker lines are thicker for users who contributed primarily to Topic 2 (“lawyers’ plight”), Topic 3 (“politically liberalism”), and Topic 15 (“professional regulation”) to facilitate their identification.

randomly. Each example that follows below is preceded by the message ID of the OP. Examples chosen purposively and randomly are denoted by “p” and “r” respectively. I include the usernames of OPs if they are among the forty super-users labeled in Figure 11. I limit the examples to seven topics: general interest, forum-related, political liberalism, nationalistic, lawyers’ plight, professional regulation, and law firms. All translations are my own.

1. General Interest: Topic 1.

Discussion threads in this topic were typically unrelated to the law and legal practice. OPs in this topic were often spawned by poems, jokes, short stories, and other efforts to stimulate small talk.

- #129368 (p, 无心): The largest in the entire archive in terms of messages, this thread consists of 531 messages from sixty-six unique users in twenty-two provinces. It was initiated by an effort to organize a virtual get-together (聚会) for people to share casual small talk, poems, and media stories on miscellaneous topics.
- #597861 (p, 天助): Containing 106 messages, this thread opens with a poem titled, “Lawyer, to love you is to want to support you.” It depicts the hard-working nature of lawyers. Users replied with humorous, often sarcastic, comments mocking the OP, some suggesting it was written by his wife.
- #495405 (p, 福州宝): “A declaration of the establishment of the ‘ACLA Hooligan Alliance.’.... Membership requirements: (1) enjoy miscellaneous small talk, (2) have thick skin or, for those who are thin-skinned, the aspiration to have thick skin, and (3) women in particular are welcome to become members.” A long series of eighty-eight replies (ten unique users) contains jokes and double entendres.
- #614384 (p): A poem triggered a long thread of 115 replies (13 unique users) containing ruminations and reflections on the ideals and realities of love. The OP is a cynical and bitter portrayal of the course of marriage from the perspective of a man increasingly henpecked by his wife.
- #570174 (p, 天助): A short story about a female designer’s personal and professional life led to 126 comments from twenty-three unique users.
- #375849 (p): A bittersweet Spring Festival (Chinese New Year) poem posted on January 26, 2004 about returning to the home

village of one's youth inspired 108 replies from thirty-three unique users.

- #619520 (r): This randomly-selected short thread was actually posted by a non-lawyer seeking legal advice. The OP claimed to have been regularly accused of being homosexual because he was 30 and unmarried. In the recent past, he physically attacked a colleague who publicly accused him of having a "special hobby" and "same-sex attraction." "I've had a few girlfriends, but never took the marriage route. . . . Where I live 30 is considered old, and if you're unmarried people will relentlessly accuse you of being homosexual, which is a far harsher insult than being called a thief." The OP asked if he could sue his accuser for slander. He received three replies.
- #434778 (r): The OP is a long joke about flatulence, followed by a single non sequitur joke about the origins of the Japanese flag.
- #514629 (r): This randomly-selected thread contains 26 replies from 21 unique users to a judge seeking career and life counseling. He or she expressed feeling stuck in a rut, trapped in a boring and meaningless job that supported only a subsistence-level quality of life. Some repliers encouraged the OP to consider becoming a lawyer.

2. *Lawyers' Plight: Topic 2.*

The topic model placed discussion threads about professional challenges and vulnerabilities, particularly with respect to clients and bar associations, in this topic category.

- #58727 (p): In the largest thread in this topic, the OP suggested establishing an official Lawyers Day. Many users (149 unique users posting 190 replies) posted humorous proposals for dates. The discussion shifted to the general topic of lawyers' weak status in society. One user wrote, "What we need is higher status in society and legal rights, not a brief happy celebration! Even if a Lawyers Day is established, we will still have difficulty collecting evidence, difficulty accessing clients in police custody, and difficulty defending our clients."
- #627442 (p, 李君友): The OP is a news story about lawyers who were successfully sued by their clients and ordered to compensate them ¥8 million (approximately US \$1 million in 2004). The discussion shifted to the more general topic of professional risks and lawyers' reputational challenges.

- #572621 (p): The OP is the famous labor lawyer, Zhou Litai (周立太), who specialized in representing migrant workers owed wages and compensation.⁹³ He asked for advice on how to collect his legal fees from clients for whom he successfully won compensation. The discussion it stimulated concerned coping strategies in a marketizing professional context characterized by weak social trust and weak legal protections afforded to lawyers.
- #376771 (r): This thread contains a discussion about the virtues of working as in-house counsel. According to users who contributed to this thread, given that law firms provided little to no support with respect to finding clients (案源), billing clients, ensuring a stable income, and providing administrative support, working as in-house counsel obviated many of the insecurities faced by lawyers.
- #141935 (r): In this inflammatory tirade against the ACLA, the OP asked, “What is the ACLA able to do aside from collecting fees and selling gowns? I mean what the hell are they? It would be better to issue Tang-style suits (唐装), for at least they look traditional.” The discussion thread then shifted to an attack against “legal workers” (法律工作者, or 法工), who allegedly started wearing black gowns to masquerade as lawyers in court in order to steal their business.

3. *Political Liberalism: Topic 3.*

Discussions in this topic concern institutional constraints (and the lack thereof) on political interference in the legal system, and political freedoms more generally. Notable activists do appear in this topic, both as users and as subjects of discussion. The well-known cases of official harassment, intimidation, physical abuse, and even prosecution of lawyers such as Gao Zhisheng (高智晟), Li Kuisheng (李奎生), Teng Biao (滕彪), Zheng Enchong (郑恩宠), and Zhang Jianzhong (张建中) were discussed on the ACLA forum (not always sympathetically, however). Several prominent lawyer activists used their real names as their usernames, such as Guo Guoting (郭国汀), Tang Jingling (唐荆陵), Yu Wensheng (余文生), Zheng Jianwei (郑建伟), Huang Liangwei (黄梁为), Pu Zhiqiang (浦志强), and Wang Chengbiao (王成标). They

⁹³ Ethan Michelson, *The Practice of Law as an Obstacle to Justice: Chinese Lawyers at Work*, 40 L. & SOC'Y REV. 1, 18–21 (2006).

participated in discussions spanning numerous topics, including political liberalism. Messages posted by Guo Guoting and Pu Zhiqiang were particularly concentrated in this topic. Peking University's He Weifang (贺卫方) was a registered user who posted one message in the entire archive. Among the eighty replies from thirty-eight unique users in fourteen provinces to an OP (#450759) about a recent speech by He Weifang was a message by He Weifang himself, which simply stated, "Just testing to see if I can post stuff." Among the issues raised in this thread—which the model classified as belonging to the political liberalism topic—was the organized crime case of Liu Yong (刘涌), the subject of an even longer thread two months earlier.

- #385063 (p): With 101 replies from twenty-seven unique users, this discussion thread was the largest in this topic. The OP was about a landmark legal victory for the Chinese legal profession. Chen Xingliang (陈兴良), a professor at Peking University, was part of an advisory group of experts who contributed to the appeal of mob boss Liu Yong's death sentence in Shenyang, Liaoning Province.⁹⁴ The death sentence was not overturned, but converted to a two-year reprieve, which in practice almost always means a life sentence in prison. The OP accused "legal elites" like Chen Xingliang of "sacrificing principle for profit" (见利忘义) and of elevating their professional and public reputations at the expense of procedural justice by unduly influencing public opinion and thereby rigging the process. Other users challenged the OP by arguing that the court's written decision, which contained, "a high probability of a confession coerced through torture," was an important step towards strengthening the protection of defendants' rights and due process. Users were split between casting Chen Xingliang as a villain and hailing him as champion of the rule of law and human rights.
- #226970 (p): The OP, incidentally written by Guo Guoting, a prominent activist lawyer reportedly now in exile in Canada, concerned the case of his client, Zheng Enchong (郑恩宠), another activist lawyer known for representing residents forcibly

⁹⁴ Pils, *supra* note 11, at 1247 n.133 (discussing the same case but mistakenly placing it in Heilongjiang Province). If this case was a legal victory, it was a short-lived one. The Supreme People's Court subsequently reinstated the original death sentence reportedly because the Minister of Public Security, Zhou Yongkang, "did not want a precedent set allowing confessions obtained by torture to be thrown out of court." John Pomfret, *Execution Reveals Party's Grip in China*, WASH. POST, Dec. 23, 2003, at A12.

evicted from their homes in demolition and relocation cases. His co-counsel was Zhang Sizhi (张思之), another notable activist lawyer whose main claim to fame is representing the Gang of Four in 1981. Zheng was prosecuted for leaking state secrets and ultimately sentenced to three years in prison. The ninety-two replies from forty-two unique users decried what they understood as a political assault against citizen's rights generally and lawyers' rights in particular. Users also expressed grave concerns for the physical safety of Guo Guoting and Zhang Sizhi and the political risks they were assuming in light of political interference in this politically sensitive case.

Although ACLA forum users clearly articulated the strong importance they attached to the protection of legal rights, they rarely used the language of "rights defense" (*weiquan*, 维权). To be sure, they did invoke the language of "rights defense." Of all 24,645 discussion threads in the archive, 630 (2.5%) contain the word *weiquan* (维权). Users who did invoke it, however, invoked it in its literal sense endorsed by the party-state: to use the law as a weapon.⁹⁵ The rarity of its use among lawyers cannot be explained by its political sensitivity, for it was not a politically sensitive word.⁹⁶ Perhaps lawyers rarely invoked this word because of its banality. Indeed, almost half of all lawyers in a large nationwide survey conducted in 2009 self-identified as *weiquan* lawyers.⁹⁷ When asked to identify the defining characteristics of *weiquan* lawyers, many respondents indicated that the word "*weiquan*" was synonymous with the word "lawyer" and was thus devoid of political significance.⁹⁸ It referred simply to the work of lawyers: defending the rights and interests of clients.⁹⁹ This should be unsurprising given that Article 2 of the Lawyers Law stipulates, "Lawyers should protect their clients' lawful rights and interests" (律师应当维护当事人合法权益).¹⁰⁰ Rarely was the term *weiquan* used to refer to a specific type of lawyer. Only five discussion

⁹⁵ See generally Mary E. Gallagher, *Mobilizing the Law in China: "Informed Disenchantment" and the Development of Legal Consciousness*, 40 L. & SOC'Y REV. 783 (2006); GALLAGHER, *supra* note 19.

⁹⁶ Pils, *supra* note 11, at 1225 (Pils acknowledges this point: the word *weiquan* was "by itself not politically 'sensitive'").

⁹⁷ Michelson & Liu, *supra* note 1, at 323.

⁹⁸ *Id.*

⁹⁹ See generally Ethan Michelson, *Many Voices in China's Legal Profession: Plural Meanings of Weiquan*, 4 CHINA LAW & SOC'Y REV. 71 (2019); BENNEY, *supra* note 25.

¹⁰⁰ Lawyers Law (律师法) (promulgated by the Standing Comm. Nat'l People's Cong., May. 15, 1996, rev'd Oct. 28, 2007, effective June. 1, 2008) art. 2, 2007 P.R.C. Laws 76 (China), available at <https://perma.cc/7ZZF-LUR4>; see also Pils, *supra* note 11, at 1225.

threads contain the term “维权人士” (rights defender) and only nine contain the term “维权律师” (rights defense lawyer), both of which were ostensibly in wide circulation at this time.¹⁰¹ *Weiquan* is simply a contraction for *weihu . . . hefa quanyi* (维护 . . . 合法权益, “to protect [someone’s] lawful rights and interests”). This full expression of *weiquan* appeared in 645 (or 2.6%) of all discussion threads (e.g., 维护当事人的合法权益, 维护自己的合法权益, 维护律师的合法权益, etc.). Of all 24,645 discussion threads, only six include the term “*weiquan* movement” (维权运动), three of which were classified as belonging to the “political liberalism” topic, including one thread about rural land expropriation (#452229).¹⁰²

4. Nationalistic: Topic 5.

No different from the majority of online political discourse on Chinese websites,¹⁰³ the ACLA forum contained a large volume of nationalistic sentiment expressing, for example, opposition to Western imperialist efforts to prevent China from realizing its historical destiny and from reviving its glorious historical legacy. Much of the discourse in this topic also concerned atrocities committed by Japanese military forces during its occupation of China and the status of Taiwan.

- #564647 (p): With 124 messages from thirty-four unique users in twelve unique provinces, the largest thread in this topic exemplifies the extent of Japan-bashing on the ACLA forum. After the OP accused another user of being an admirer of and apologist for Japan (“if you don’t hate the Japanese, you can’t call yourself Chinese”), users piled on and escalated the rhetoric to a broader attack against President George W. Bush, Prime Minister Junichiro Koizumi, and what was expressed as the false facade of Western democracy. Some users, however, pushed back against anti-Japanese sentiments by arguing that it serves the interests of

¹⁰¹ See generally Pils, *supra* note 11.

¹⁰² The OP of this thread was a report on a workshop devoted to the issue of village elections in the context of rural “land grabs.” Incidentally, the workshop was organized at the behest of Zhao Yan (赵岩), who at the time was employed by *The New York Times*, and who was subsequently charged for leaking state secrets (only five months after this message was posted). He was found guilty of fraud, not of leaking state secrets, and sentenced to three years in prison. Jim Yardley & Joseph Kahn, *Top Charge Dropped, but China Gives Times Researcher 3 Years*, N.Y. TIMES, Aug. 25, 2006, at A1.

¹⁰³ See generally ONLINE CHINESE NATIONALISM AND CHINA’S BILATERAL RELATIONS, *supra* note 32; XU WU, *supra* note 32.

the state by distracting attention away from China's domestic social and political problems.

- #477449 (p, 骆伟雄): Following an OP on the hypocrisy of US human rights in light of Abu Ghraib torture and prisoner abuse, some users pushed back. Whereas, according to some users, China will never admit to its pervasive use of torture to extract confessions, the US remains an exemplar of human rights because it can confront its demons openly. One user's response to this argument: "as if everything the stinky dog shit US does is so great" (真是美国的臭狗屎都是好的). To some users' critique that the OP was overly influenced by Chinese media propaganda, the OP responded by arguing that the US case was about state-sponsored torture, whereas in China the problem of torture was about individual bad apples.
- #534446 (p, 人种): The OP asked users to "[i]magine, hypothetically, if another Cultural Revolution were launched to deal with rampant corruption. Some people say that if we simply executed every government official at the county level and higher, the result would be the same as if they had been tried in court." The argument was that a Maoist-style purge would be more efficient and effective than trials protecting procedural rights and due process. Many but not all users supported this position. Whereas some users argued that the Cultural Revolution was the true manifestation of a people's democracy and invoked the famous song, "Chairman Mao, the Sun in Our Hearts," others pushed back by changing the title to "the sun in the hearts of some of us."
- #556325 (p, jlg): This thread is about James Soong and Lien Chan—high-level Taiwanese politicians—in the context of the 2004 riots following China's defeat to Japan in a soccer match. The OP posed a question—"why can't Chinese people admit defeat?"—and answered it with reference to China's historical humiliations and concomitant inferiority complex. The connection to Taiwan is the puzzle of Taiwan's unwillingness to admit defeat and reunite with the Mainland. In response to the motivating question of why Chinese people had trouble accepting China's loss of the soccer match to Japan, one user wrote: "See how you feel after your wife gets abducted, your daughter is sold to Thailand as a slave, and your home is demolished without compensation, and then your neighbor who murdered your father gloats in front of your face every day. If you can remain rational

then hats off to you!” The words “nationalism” (民族主义) and “patriotic” (爱国) appear throughout this thread. But not all users accepted the generally anti-Japanese sentiment. In the words of one user, “Since China became a nuclear power, Japanese tanks haven’t rolled into Tiananmen Square. It was our own tanks. We need so solve our own problems first before solving other people’s problems.”

- #463702, (p, 海子律): This discussion thread was devoted to high praise for Premier Zhu Rongji after his final appearance on the front stage of the Great Hall of the People, and reflects a cult of personality enjoyed by some of China’s top political leaders.
- #463566 (p, 鲁典): The OP is a news report of a Japanese police investigation following tips from Japanese netizens concerning a message posted on a discussion board about bombing the Chinese embassy in Japan. This report sparked all kinds of outrage in the thread, including suggestions that China make a preemptive military strike against Japan and declarations about the inevitability of China’s rise to surpass Japan. One user responded to the suggestion of a military strike by writing, “Don’t boast about bombing Japan when we still haven’t dealt with Taiwan.” Another user was impressed that the apparent bomb threat attracted the attention of the Japanese police given that there would be nothing unusual about a message posted on a Chinese message board expressing an intent to kill Japanese people. One user (a super-user depicted in Figure 11, 吴秀夫) wrote: “It’s a pity people who shout slogans at the top of their lungs about flattening the island of Japan with bombs, killing all Japanese men, and raping Japanese women don’t know how to put their words into action.” The discussion then shifted to the US bombing of the Chinese embassy in Belgrade.
- #531202 (p, 吴秀夫): In the context of the trial of Saddam Hussein, the OP posted a news story about ordinary Iraqi citizens proclaiming undying loyalty to Saddam as their rightful leader. Another user quoted, from a separate and unattributed source, an Iraqi citizen declaring Saddam a courageous hero for standing up to the US. In their replies, users expressed support for Saddam as the legitimate, duly elected president of Iraq, their belief that his trial was illegitimate, the idea that Chinese lawyers should lend their legal support and assistance to his cause, and their opinion

that the US-led coalition forces should leave Iraq. The discussion shifted to US atrocities in occupied Iraq.

5. *Forum-Related: Topic 6.*

This topic contains discussions that are primarily requests and suggestions related to the administration of the ACLA forum. Several threads discussed disputes between users.

- #420428 (p): The OP that launched this thread of 387 messages from sixty-eight unique users was a list of prominent users (followed by many replies from users wondering why they were not included on it).
- #502998 (p, 听雨轩): The OP requested that the username of another user (宋都狂犬, “Song Capital Rabies”) be deleted on the grounds that it is historically offensive to the city of Kaifeng, Henan Province, one of the capitals of the Song Dynasty.
- #547778 (p, jlg): The female OP reported sexual harassment via email, and asked, “Which hooligan did this?” She continued with the following: “The only place I’ve ever used a female avatar is on the ACLA forum; nowhere else have I used one, and no one knows my gender.” She reported receiving a sexually suggestive solicitation by email from a self-reported “handsome young man in Shanghai.” The OP received eighty replies from fifty unique users.

6. *Professional Regulation: Topic 15.*

This topic drips with vitriol directed against “legal workers” and the government agencies that allow them to compete with lawyers.¹⁰⁴

- #528656 (p, jlg): In the largest discussion thread in this topic (120 messages posted by forty-one unique users in eleven unique provinces), users debated who should count as a lawyer, and how the legal services market should be better regulated.
- #472047 (p, 海子律): The OP of this thread illustrated the challenge legal workers posed to lawyers with a joke about a former judge currently working as a legal worker who used his courthouse connections to out-compete lawyers. Other users

¹⁰⁴ See Sida Liu, *Lawyers, State Officials and Significant Others: Symbiotic Exchange in the Chinese Legal Services Market*, 206 CHINA Q. 276 (2011).

replied with claims about how legal workers were a key source of judicial corruption and thereby inflicted damage to lawyers' public image.

- #230626 (p): The OP was a call to eliminate the “law offices” of legal workers that were allegedly proliferating in Chinese cities, on major boulevards and small alleys alike, particularly around courthouses. In a reply, one user wrote, “Basic-level legal workers have become a disaster for the legal system and a disaster for lawyers! They are a disaster to litigation!”
- #5099 (r): The OP posted a query about how to obtain a license to practice securities law. One user replied, “I heard the Ministry of Justice plans to abolish the securities license and open up this part of the legal services market.” Another user replied, “The exam was halted. Expect an announcement from the Ministry of Justice.” A third user replied: “Restricting the market with a special license constitutes unfair competition. The license will likely be abolished.”
- #554486 (r): The OP posted a request for help representing two lawyers in Luoyang, Henan Province, Li Subin (李苏滨) and Li Wuxi (李午汜), in their administrative compensation lawsuit against the Luoyang Municipal Bureau of Justice, which had refused to renew their lawyers licenses because of old allegations that they had collected legal fees directly from clients (i.e., failed to report income, embezzled from their law firms). In contrast to the widespread sympathy and support reported by Liu and Halliday in response to lawyers' efforts to mobilize support in similar cases, many lawyers posted replies to this thread scornfully criticizing the two lawyers for shamelessly seeking free legal representation. Some of the discussion centered on the scope of legal aid, and whether the lawyers would qualify. They ultimately won their case; the Luoyang Municipal Bureau of Justice was ordered to compensate the lawyers for lost earnings.
- #548145 (r): This thread contains four replies to an OP who listed reasons for his or her opposition to legal workers.

7. *Law Firms: Topic 17.*

Discussions in this topic concern lawyer compensation, law firm organization, branch offices, and related topics.

- #38227 (p): With 231 messages from 218 unique users in twenty-eight unique provinces, this is the largest thread in the entire archive in terms of both unique users and unique provinces. (Only Hainan, Gansu, and Qinghai were excluded, although these provinces may have been represented by users who did not disclose their geographical locations.) The OP requested the detailed contact information of users for purposes of compiling an “ACLA Forum Netizen Address Book.”
- #97750 (p, 李君友): Similar to the above, this OP requested the contact information of firms wishing to consider ways to cooperate professionally. This was an effort to build a nationwide collaborative practice network. It received seventy-five replies from fifty-eight unique users in eighteen unique provinces.
- #71554 (p): An OP about lawyer internship compensation standards spawned fifty-four replies from twenty-seven unique users.
- #447430 (p, 人种): The OP shared a news story about two veteran China Central Television journalists who were disciplined for physically fighting each other while covering a signing ceremony at the Grand Hyatt Hotel in Beijing for a US-China airspace cooperation treaty. The US delegation included members of Congress. This discussion thread was likely classified in the “law firms” topic because it contains several instances of the words “journalist” and “news.”
- #473849 (p, 骆伟雄): The OP, based in Guangxi Province, was considering opening a Beijing branch office of his law firm.
- #11546 (r): Although the OP concerned a criminal case in rural Guangxi Province, the discussion it spawned was squarely about law firms. The OP shared a news story about a female schoolteacher who was hacked to death in public by her husband. When the chief of the local Bureau of Justice arrived at the crime scene the victim was still alive. He not only failed to attempt to save her, but also prevented others from trying to do so. He was subsequently charged criminally for professional negligence, but the local procuratorate decided not to prosecute. Inexplicably, the discussion, which includes thirty-five replies from twenty-eight unique users, pivoted to lawyer income in general and a comparison of salaries by law firm and city in particular after a repplier asked whether a lawyer in Beijing can earn ¥100,000 per year (about US \$12,000 in 2004).

- #411279 (r): The OP asked whether lawyer interns need licenses.
- #578353 (r): A call to prospective partners interested in establishing a new firm.
- #623679 (r): “I’m starving! Let’s go eat! I’m so unfortunate to have been called in to work over the weekend.”

V. DISCUSSION AND CONCLUSIONS

Findings I presented from an analysis of a large corpus of discussions posted—mostly in 2004—to the largest online community of lawyers that ever existed in China comports with and supports much of what we know from the existing literature about Chinese lawyers. Much of what I have reported is unsurprising. We already knew that lawyers in the mid-2000s were vocal in their discontent with their professional conditions. We knew a lot about their efforts to mobilize in support of better and more strongly enforced legal rights for themselves and their clients. And we knew that lawyers framed their complaints and hopes in the language of political liberalism. Lawyers’ *diagnostic* understandings of—and their *prescriptive* remedies to—their collective plight were often expressed in terms of political liberalism. As they saw it, weak rule of law and basic legal freedoms coupled with routine political interference stymied their professional interests (and those of their clients). As they also saw it, mobilizing and directing their collective efforts towards strengthening the rule of law and constraining the obstructionist tendencies of the state offered the promise of better professional conditions.

Some of what I have reported, however, may be more surprising. Prior to this study, we lacked a more macroscopic view of the broader landscape of lawyer discourse of all stripes in which their narrowly politically liberal discourse was situated. We now appreciate that Chinese lawyers in 2004 embraced both politically liberal and politically illiberal values. The title of this article includes “heyday of political activism.” Previous research tells us the ACLA forum was an important vehicle for lawyer mobilization: “Over and over again lawyers called upon each other [in messages posted to the ACLA forum] to take political action.”¹⁰⁵ Findings I presented show that this is true in more ways than one. In addition to calling for mobilization in support of political liberalism, lawyers also called for mobilization in support of pro-regime, illiberal, and nationalistic causes.

¹⁰⁵ LIU & HALLIDAY, *supra* note 4, at 149.

Although it was palpable and unambiguous, lawyers' politically liberal discourse occupied only a relatively small part of the ACLA forum's content. Most of the forum's content consisted of general interest and practice-oriented discussions. Findings from a more comprehensive analysis of lawyers' online discourse fail to support Liu and Halliday's assertion that messages related to "criminal procedure law and the work of defense lawyers had been among the central topics and most heavily trafficked discussion sites since the beginning of the forum."¹⁰⁶ Moreover, a huge share of the corpus of discussions I analyzed was concentrated among a relatively small share of super-users. All discussions on all topics, both political and practice-oriented alike, were dominated by a relatively small handful of extraordinarily prolific users (including Sida Liu himself) who contributed the vast majority of the ACLA forum's content.

Finally, if the ACLA forum was a hotbed of political activism and, as Liu and Halliday argue, the birthplace of a "liberal moment" in China,¹⁰⁷ few of its participants knew about it. If *weiquan* lawyers were an emerging force in China, as scholars have led us to believe,¹⁰⁸ they seem to have avoided the ACLA forum, and those who did frequent the ACLA forum seemed to avoid discussing them. Perhaps they were an emerging force but simply absent from discussions on the ACLA forum. Alternatively, perhaps they were not an emerging force. A third but implausible possibility is that they did participate in discussions on the ACLA forum but obscured their activist identities.

To sum up, there is no doubt that Chinese lawyers in 2004 were inspired by the ideals of political liberalism to advance and support political activism. My research helps put this segment of the Chinese legal profession in perspective. At the same time, however, the results of my analysis cast grave doubts on claims that Chinese lawyers have, to any meaningful degree, pursued political activism under the mantle of *weiquan*, or "rights defense." To be sure, the finding that *weiquan* discourse was only a marginal part of the broader landscape of lawyers' discourse in 2004 no way diminishes the significance and commitment of the handful of activists who disproportionately dominate academic and journalistic representations of China's legal profession.¹⁰⁹ However, when we talk about what lawyers talk about, nuance, perspective, and empirical rigor should be the order of the day.

¹⁰⁶ *Id.* at 14.

¹⁰⁷ See generally Halliday & Liu, *supra* note 5; LIU & HALLIDAY, *supra* note 4, at 148.

¹⁰⁸ See generally Fu & Cullen, *supra* note 13; PILS, *supra* note 9.

¹⁰⁹ See generally Stern, *supra* note 10.

APPENDIX

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
1	1	now	现在	2	5
1	2	know	知道	1	
1	3	already	已经	2	5
1	4	should	应该	2	6
1	5	problem	问题	14	2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 16, 20
1	6	friend	朋友	1	
1	7	child	孩子	1	
1	8	life	生活	1	
1	9	time	时间	6	2, 12, 13, 16, 20
1	10	somewhat	一下	2	6
1	11	feel	觉得	1	
1	12	see	看到	1	
1	13	matter	事情	1	
1	14	woman	女人	1	
1	15	man	男人	1	
1	16	like	喜欢	1	
1	17	at the time	当时	1	
1	18	certainly	一定	1	
1	19	really	真的	1	
1	20	hope	希望	2	6
1	21	thing	东西	1	
1	22	first	第一	2	5
1	23	discover	发现	1	
1	24	after	以后	1	
1	25	today	今天	1	
1	26	inform	告诉	1	
1	27	place	地方	1	
1	28	need	需要	2	15
1	29	father	父亲	1	
1	30	sense	感觉	1	
2	1	lawyer	律师	2	6
2	2	firm	事务所	1	
2	3	judge	法官	2	11
2	4	practice	执业	2	15
2	5	fee	收费	1	
2	6	litigant	当事人	4	4, 11, 16
2	7	profession	职业	3	12, 15
2	8	judicial	司法	4	11, 12, 15
2	9	association	协会	1	
2	10	case	案件	4	7, 11, 16
2	11	China (1)	中国	4	3, 5, 6
2	12	industry	行业	2	15
2	13	collect	收取	1	
2	14	relationship	关系	3	8, 13
2	15	lawsuit	官司	1	
2	16	cost	费用	3	4, 13

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
2	17	maintain	维护	1	
2	18	Ministry of Justice	司法部	2	12
2	19	work	工作	6	12, 13, 15, 17, 19
2	20	standard	标准	2	9
2	21	rights and interests	权益	1	
2	22	director	主任	2	17
2	23	problem	问题	14	1, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 16, 20
2	24	business	业务	2	15
2	25	handle	办理	3	4, 18
2	26	courtroom	法庭	3	14, 16
2	27	stage	阶段	1	
2	28	society	社会	4	3, 13, 15
2	29	environment	环境	1	
2	30	time	时间	6	1, 12, 13, 16, 20
3	1	society	社会	4	2, 13, 15
3	2	system (1)	制度	1	
3	3	nation	国家	6	5, 7, 12, 19, 20
3	4	law	法律	8	4, 8, 9, 11, 12, 15, 18
3	5	problem	问题	14	1, 2, 4, 5, 6, 8, 9, 10, 11, 12, 13, 16, 20
3	6	government	政府	4	5, 6, 19
3	7	constitution	宪法	1	
3	8	freedom	自由	1	
3	9	representative	代表	1	
3	10	authority	权力	1	
3	11	China (1)	中国	4	2, 5, 6
3	12	right	权利	2	8
3	13	democracy	民主	1	
3	14	economy	经济	5	5, 13, 15, 17
3	15	politics	政治	2	5
3	16	peasant	农民	1	
3	17	reform	改革	1	
3	18	People's Congress	人大	1	
3	19	people	人民	6	5, 11, 14, 18, 19
3	20	interest	利益	2	8
3	21	development	发展	3	5, 17
3	22	citizen	公民	1	
3	23	legislation	立法	1	
3	24	resolve	解决	1	
3	25	leader	领导	1	
3	26	rule of law	法治	1	
3	27	think	认为	7	5, 6, 7, 8, 14, 16
3	28	important	重要	1	
3	29	become	成为	2	5
3	30	exist	存在	3	8, 16
4	1	contract	合同	2	13
4	2	agree	约定	1	
4	3	agreement	协议	1	
4	4	housing	房屋	1	

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
4	5	implement	履行	1	
4	6	mutual	双方	2	8
4	7	stipulation	规定	9	7, 8, 9, 10, 11, 13, 18, 19
4	8	breach of contract	违约	1	
4	9	sign	签订	1	
4	10	demand	要求	4	12, 14, 16
4	11	pay	支付	2	13
4	12	handle	办理	3	2, 18
4	13	litigant	当事人	4	2, 11, 16
4	14	whether	是否	6	7, 8, 10, 13, 16
4	15	undertake	承担	3	9, 10
4	16	invalid	无效	1	
4	17	effective	有效	1	
4	18	problem	问题	14	1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 13, 16, 20
4	19	engineering	工程	2	19
4	20	remove	解除	2	13
4	21	liability	责任	4	7, 9, 10
4	22	duty	义务	2	8
4	23	contract law	合同法	1	
4	24	losses	损失	2	9
4	25	buy and sell	买卖	1	
4	26	provision	条款	1	
4	27	law	法律	8	3, 8, 9, 11, 12, 15, 18
4	28	ought	应当	7	8, 9, 11, 13, 18, 19
4	29	deadline	期限	1	
4	30	cost	费用	3	2, 13
5	1	China (1)	中国	4	2, 3, 6
5	2	U.S.A.	美国	1	
5	3	nation	国家	6	3, 7, 12, 19, 20
5	4	population	人口	1	
5	5	Japan	日本	1	
5	6	nationality	民族	1	
5	7	Taiwan	台湾	1	
5	8	world	世界	1	
5	9	culture	文化	1	
5	10	history	历史	1	
5	11	Chinese	中国人	1	
5	12	economy	经济	5	3, 13, 15, 17
5	13	people	人民	6	3, 11, 14, 18, 19
5	14	development	发展	3	3, 17
5	15	government	政府	4	3, 6, 19
5	16	problem	问题	14	1, 2, 3, 4, 6, 8, 9, 10, 11, 12, 13, 16, 20
5	17	plan	计划	1	
5	18	war	战争	1	
5	19	international	国际	2	20
5	20	now	现在	2	1
5	21	think	认为	7	3, 6, 7, 8, 14, 16
5	22	politics	政治	2	3

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
5	23	-ism	主义	1	
5	24	policy	政策	1	
5	25	China (2)	中华	2	11
5	26	become	成为	2	3
5	27	element	分子	1	
5	28	territory	地区	2	17
5	29	first	第一	2	1
5	30	already	已经	2	1
6	1	community	社区	1	
6	2	Netizen	网友	1	
6	3	forum	论坛	1	
6	4	Mr.	先生	1	
6	5	forum moderator	斑竹	1	
6	6	essay	文章	1	
6	7	opinion	意见	1	
6	8	discuss	讨论	1	
6	9	support	支持	2	14
6	10	post	帖子	1	
6	11	lawyer	律师	2	2
6	12	China (1)	中国	4	2, 3, 5
6	13	government	政府	4	3, 5, 19
6	14	network	网络	2	20
6	15	suggestion	建议	1	
6	16	website	网站	2	20
6	17	publish	发表	1	
6	18	manage	管理	4	10, 18, 19
6	19	register (1)	注册	2	20
6	20	problem	问题	14	1, 2, 3, 4, 5, 8, 9, 10, 11, 12, 13, 16, 20
6	21	rule	规则	1	
6	22	punish	处罚	3	7, 18
6	23	hope	希望	2	1
6	24	think	认为	7	3, 5, 7, 8, 14, 16
6	25	thank you	谢谢	1	
6	26	should	应该	2	1
6	27	reality	现实	1	
6	28	disclose	公开	1	
6	29	description	说明	1	
6	30	somewhat	一下	2	1
7	1	crime	犯罪	1	
7	2	criminal	刑事	1	
7	3	behavior	行为	4	8, 14, 18
7	4	agency	机关	2	18
7	5	public security	公安	1	
7	6	constitute	构成	1	
7	7	defendant (2)	被告人	1	
7	8	personnel	人员	4	15, 17, 19
7	9	stipulation	规定	9	4, 8, 9, 10, 11, 13, 18, 19
7	10	defend	辩护	1	

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
7	11	unlawful	非法	1	
7	12	criminal law	刑法	1	
7	13	procuratorate	检察院	2	11
7	14	severe	严重	1	
7	15	punish	处罚	3	6, 18
7	16	suspect	嫌疑人	1	
7	17	police	警察	1	
7	18	case	案件	4	2, 11, 16
7	19	plot	情节	1	
7	20	public security bureau	公安局	1	
7	21	whether	是否	6	4, 8, 10, 13, 16
7	22	nation	国家	6	3, 5, 12, 19, 20
7	23	prison sentence	徒刑	1	
7	24	think	认为	7	3, 5, 6, 8, 14, 16
7	25	carry out	进行	5	10, 16, 19, 20
7	26	below	以下	1	
7	27	cause	造成	2	9
7	28	liability	责任	4	4, 9, 10
7	29	situation	情况	4	8, 10, 16
7	30	occur	发生	3	9, 13
8	1	behavior	行为	4	7, 14, 18
8	2	law	法律	8	3, 4, 9, 11, 12, 15, 18
8	3	property	财产	1	
8	4	relationship	关系	3	2, 13
8	5	stipulation	规定	9	4, 7, 9, 10, 11, 13, 18, 19
8	6	right	权利	2	3
8	7	marriage	婚姻	1	
8	8	joint	共同	1	
8	9	think	认为	7	3, 5, 6, 7, 14, 16
8	10	divorce	离婚	1	
8	11	problem	问题	14	1, 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 16, 20
8	12	whether	是否	6	4, 7, 10, 13, 16
8	13	duty	义务	2	4
8	14	principle	原则	1	
8	15	marital	夫妻	1	
8	16	protect	保护	2	20
8	17	obtain	取得	1	
8	18	third	第三	2	9
8	19	exist	存在	3	3, 16
8	20	ought	应当	7	4, 9, 11, 13, 18, 19
8	21	belong	属于	1	
8	22	debt	债务	2	10
8	23	mutual	双方	2	4
8	24	debt claim	债权	1	
8	25	marry	结婚	1	
8	26	amongst	之间	1	
8	27	applicable	适用	2	11
8	28	interest	利益	2	3

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
8	29	register (2)	登记	3	10, 18
8	30	situation	情况	4	7, 10, 16
9	1	liability	责任	4	4, 7, 10
9	2	compensation	赔偿	1	
9	3	traffic	交通	1	
9	4	accident	事故	1	
9	5	undertake	承担	3	4, 10
9	6	harm	损害	1	
9	7	insurance	保险	2	13
9	8	motor vehicle	机动车	1	
9	9	road	道路	1	
9	10	stipulation	规定	9	4, 7, 8, 10, 11, 13, 18, 19
9	11	cause	造成	2	7
9	12	occur	发生	3	7, 13
9	13	losses	损失	2	4
9	14	vehicle	车辆	1	
9	15	ought	应当	7	4, 8, 11, 13, 18, 19
9	16	death	死亡	2	16
9	17	tort	侵权	2	20
9	18	bodily	人身	1	
9	19	spirit	精神	1	
9	20	fault	过错	1	
9	21	third	第三	2	8
9	22	standard	标准	2	2
9	23	victim	受害人	1	
9	24	coerce	强制	1	
9	25	civil	民事	3	11, 14
9	26	law	法律	8	3, 4, 8, 11, 12, 15, 18
9	27	problem	问题	14	1, 2, 3, 4, 5, 6, 8, 10, 11, 12, 13, 16, 20
9	28	explanation	解释	2	11
9	29	security	安全	2	19
9	30	deal with	处理	1	
10	1	company	公司	1	
10	2	enterprise	企业	4	13, 19, 20
10	3	bank	银行	1	
10	4	land	土地	1	
10	5	shareholder	股东	1	
10	6	invest	投资	1	
10	7	assets	资产	1	
10	8	operate	经营	3	19, 20
10	9	limited	有限	1	
10	10	risk	风险	1	
10	11	funds	资金	1	
10	12	liability	责任	4	4, 7, 9
10	13	state-owner	国有	1	
10	14	problem	问题	14	1, 2, 3, 4, 5, 6, 8, 9, 11, 12, 13, 16, 20
10	15	transfer	转让	1	
10	16	commercial	商业	1	

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
10	17	legal person	法人	1	
10	18	register (2)	登记	3	8, 18
10	19	loan	贷款	1	
10	20	transaction	交易	1	
10	21	stipulation	规定	9	4, 7, 8, 9, 11, 13, 18, 19
10	22	carry out	进行	5	7, 16, 19, 20
10	23	finance	金融	1	
10	24	set up	设立	1	
10	25	undertake	承担	3	4, 9
10	26	manage	管理	4	6, 18, 19
10	27	whether	是否	6	4, 7, 8, 13, 16
10	28	situation	情况	4	7, 8, 16
10	29	collective	集体	1	
10	30	debt	债务	2	8
11	1	court	法院	2	14
11	2	people	人民	6	3, 5, 14, 18, 19
11	3	case	案件	4	2, 7, 16
11	4	judge	法官	2	2
11	5	enforce	执行	1	
11	6	litigant	当事人	4	2, 4, 16
11	7	stipulation	规定	9	4, 7, 8, 9, 10, 13, 18, 19
11	8	judicial	司法	4	2, 12, 15
11	9	law	法律	8	3, 4, 8, 9, 12, 15, 18
11	10	highest	最高	1	
11	11	litigation	诉讼	3	14, 15
11	12	hearing	审理	2	14
11	13	verdict	判决	2	14
11	14	procedure	程序	2	18
11	15	civil	民事	3	9, 14
11	16	explanation	解释	2	9
11	17	trial	审判	1	
11	18	apply	申请	3	13, 18
11	19	applicable	适用	2	8
11	20	problem	问题	14	1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 13, 16, 20
11	21	procedure law	诉讼法	1	
11	22	ought	应当	7	4, 8, 9, 13, 18, 19
11	23	dispute (1)	纠纷	1	
11	24	accept	受理	2	18
11	25	ruling	裁定	1	
11	26	republic	共和国	1	
11	27	file a lawsuit	起诉	2	14
11	28	make	作出	2	18
11	29	China (2)	中华	2	5
11	30	procuratorate	检察院	2	7
12	1	examination	考试	1	
12	2	school	学校	1	
12	3	student	学生	1	
12	4	university	大学	1	

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
12	5	judicial	司法	4	2, 11, 15
12	6	education	教育	1	
12	7	teacher	老师	1	
12	8	specialty	专业	2	15
12	9	participate	参加	1	
12	10	study	学习	1	
12	11	professor	教授	1	
12	12	legal studies	法学	1	
12	13	law	法律	8	3, 4, 8, 9, 11, 15, 18
12	14	Ministry of Justice	司法部	2	2
12	15	nation	国家	6	3, 5, 7, 19, 20
12	16	work	工作	6	2, 13, 15, 17, 19
12	17	this year	今年	2	17
12	18	knowledge	知识	2	20
12	19	qualification	资格	2	15
12	20	problem	问题	14	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 13, 16, 20
12	21	certificate	证书	1	
12	22	research	研究	1	
12	23	time	时间	6	1, 2, 13, 16, 20
12	24	promulgate	公布	1	
12	25	profession	职业	3	2, 15
12	26	theory	理论	1	
12	27	demand	要求	4	4, 14, 16
12	28	foundation	基础	1	
12	29	ability	能力	2	15
12	30	specialist	专家	2	16
13	1	labor	劳动	1	
13	2	unit	单位	2	19
13	3	insurance	保险	2	9
13	4	work injury	工伤	1	
13	5	employee	职工	1	
13	6	wages	工资	1	
13	7	arbitration	仲裁	1	
13	8	work	工作	6	2, 12, 15, 17, 19
13	9	laborer	劳动者	1	
13	10	hire	用人	1	
13	11	stipulation	规定	9	4, 7, 8, 9, 10, 11, 18, 19
13	12	enterprise	企业	4	10, 19, 20
13	13	dispute (2)	争议	1	
13	14	relationship	关系	3	2, 8
13	15	pay	支付	2	4
13	16	affirm	认定	3	14, 16
13	17	guarantee	保障	1	
13	18	remove	解除	2	4
13	19	contract	合同	2	4
13	20	problem	问题	14	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 16, 20
13	21	regulation (1)	条例	1	
13	22	apply	申请	3	11, 18

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
13	23	ought	应当	7	4, 8, 9, 11, 18, 19
13	24	society	社会	4	2, 3, 15
13	25	department	部门	3	18, 19
13	26	occur	发生	3	7, 9
13	27	economy	经济	5	3, 5, 15, 17
13	28	whether	是否	6	4, 7, 8, 10, 16
13	29	time	时间	6	1, 2, 12, 16, 20
13	30	cost	费用	3	2, 4
14	1	defendant (1)	被告	1	
14	2	plaintiff	原告	1	
14	3	litigation	诉讼	3	11, 15
14	4	court	法院	2	11
14	5	verdict	判决	2	11
14	6	appellant	上诉人	1	
14	7	file a lawsuit	起诉	2	11
14	8	request	请求	1	
14	9	fact	事实	2	16
14	10	demand	要求	4	4, 12, 16
14	11	evidence	证据	2	16
14	12	time limit	时效	1	
14	13	court decision	判决书	1	
14	14	initiate/file	提起	1	
14	15	agent	代理人	2	15
14	16	civil	民事	3	9, 11
14	17	think	认为	7	3, 5, 6, 7, 8, 16
14	18	behavior	行为	4	7, 8, 18
14	19	raise	提出	2	16
14	20	prove	证明	2	16
14	21	appeal	上诉	1	
14	22	affirm	认定	3	13, 16
14	23	hearing	审理	2	11
14	24	people	人民	6	3, 5, 11, 18, 19
14	25	provide	提供	4	15, 16, 20
14	26	reason	理由	1	
14	27	support	支持	2	6
14	28	courtroom	法庭	3	2, 16
14	29	claim	主张	1	
14	30	statutory	法定	2	18
15	1	law	法律	8	3, 4, 8, 9, 11, 12, 18
15	2	service	服务	2	20
15	3	represent	代理	1	
15	4	business	业务	2	2
15	5	agent	代理人	2	14
15	6	qualification	资格	2	12
15	7	provide	提供	4	14, 16, 20
15	8	market	市场	3	17, 20
15	9	engage in	从事	1	
15	10	seek advice	咨询	1	

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
15	11	entrust	委托	1	
15	12	worker	工作者	1	
15	13	practice	执业	2	2
15	14	specialty	专业	2	12
15	15	litigation	诉讼	3	11, 14
15	16	work	工作	6	2, 12, 13, 17, 19
15	17	standardized	规范	1	
15	18	possess	具有	1	
15	19	personnel	人员	4	7, 17, 19
15	20	level	水平	1	
15	21	need	需要	2	1
15	22	society	社会	4	2, 3, 13
15	23	activity	活动	2	19
15	24	profession	职业	3	2, 12
15	25	judicial	司法	4	2, 11, 12
15	26	economy	经济	5	3, 5, 13, 17
15	27	competition	竞争	1	
15	28	scope	范围	1	
15	29	industry	行业	2	2
15	30	ability	能力	2	12
16	1	evidence	证据	2	14
16	2	appraise	鉴定	1	
16	3	prove	证明	2	14
16	4	fact	事实	2	14
16	5	hospital	医院	1	
16	6	carry out	进行	5	7, 10, 19, 20
16	7	investigate	调查	2	17
16	8	medical	医疗	1	
16	9	courtroom	法庭	3	2, 14
16	10	demolition/relocation	拆迁	1	
16	11	submit evidence	举证	1	
16	12	affirm	认定	3	13, 14
16	13	witness	证人	1	
16	14	conclusion	结论	1	
16	15	situation	情况	4	7, 8, 10
16	16	provide	提供	4	14, 15, 20
16	17	material	材料	1	
16	18	time	时间	6	1, 2, 12, 13, 20
16	19	demand	要求	4	4, 12, 14
16	20	on site	现场	1	
16	21	exist	存在	3	3, 8
16	22	case	案件	4	2, 7, 11
16	23	problem	问题	14	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 20
16	24	raise	提出	2	14
16	25	whether	是否	6	4, 7, 8, 10, 13
16	26	think	认为	7	3, 5, 6, 7, 8, 14
16	27	process	过程	1	
16	28	specialist	专家	2	12

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
16	29	death	死亡	2	9
16	30	litigant	当事人	4	2, 4, 11
17	1	Beijing	北京	1	
17	2	journalist	记者	1	
17	3	Shanghai	上海	1	
17	4	work	工作	6	2, 12, 13, 15, 19
17	5	news	新闻	1	
17	6	cooperate	合作	1	
17	7	income	收入	1	
17	8	media	媒体	1	
17	9	intern	实习	1	
17	10	city	城市	1	
17	11	report (2)	报道	1	
17	12	phone	电话	1	
17	13	center	中心	1	
17	14	resident	居民	1	
17	15	Hong Kong	香港	1	
17	16	Shenzhen	深圳	1	
17	17	connect	联系	1	
17	18	personnel	人员	4	7, 15, 19
17	19	express	表示	1	
17	20	current	目前	1	
17	21	investigate	调查	2	16
17	22	development	发展	3	3, 5
17	23	director	主任	2	2
17	24	this year	今年	2	12
17	25	territory	地区	2	5
17	26	price	价格	1	
17	27	introduce	介绍	1	
17	28	economy	经济	5	3, 5, 13, 15
17	29	market	市场	3	15, 20
17	30	group	集团	1	
18	1	administrative	行政	1	
18	2	agency	机关	2	7
18	3	stipulation	规定	9	4, 7, 8, 9, 10, 11, 13, 19
18	4	ought	应当	7	4, 8, 9, 11, 13, 19
18	5	apply	申请	3	11, 13
18	6	register (2)	登记	3	8, 10
18	7	behavior	行为	4	7, 8, 14
18	8	decision	决定	1	
18	9	by law	依法	1	
18	10	make	作出	2	11
18	11	illegal	违法	1	
18	12	department	部门	3	13, 19
18	13	license	许可	1	
18	14	punish	处罚	3	6, 7
18	15	applicant	申请人	1	
18	16	law	法律	8	3, 4, 8, 9, 11, 12, 15

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
18	17	investigation	审查	1	
18	18	regulation (2)	法规	1	
18	19	handle	办理	3	2, 4
18	20	judicial review	复议	1	
18	21	item	事项	1	
18	22	violate	违反	1	
18	23	punishment	处分	1	
18	24	accept	受理	2	11
18	25	statutory	法定	2	14
18	26	manage	管理	4	6, 10, 19
18	27	people	人民	6	3, 5, 11, 14, 19
18	28	as follows	下列	1	
18	29	procedure	程序	2	11
18	30	cancel	撤销	1	
19	1	manage	管理	4	6, 10, 18
19	2	organ	机构	1	
19	3	department	部门	3	13, 18
19	4	stipulation	规定	9	4, 7, 8, 9, 10, 11, 13, 18
19	5	personnel	人员	4	7, 15, 17
19	6	people	人民	6	3, 5, 11, 14, 18
19	7	ought	应当	7	4, 8, 9, 11, 13, 18
19	8	work	工作	6	2, 12, 13, 15, 17
19	9	government	政府	4	3, 5, 6
19	10	unit	单位	2	13
19	11	measure	办法	1	
19	12	nation	国家	6	3, 5, 7, 12, 20
19	13	carry out	进行	5	7, 10, 16, 20
19	14	party in charge	主管	1	
19	15	construct	建设	1	
19	16	organization	组织	1	
19	17	security	安全	2	9
19	18	supervise	监督	1	
19	19	project	项目	1	
19	20	document	文件	1	
19	21	activity	活动	2	15
19	22	State Council	国务院	1	
19	23	enterprise	企业	4	10, 13, 20
19	24	approve	批准	1	
19	25	manufacture	生产	1	
19	26	committee	委员会	1	
19	27	be in charge of	负责	1	
19	28	operate	经营	3	10, 20
19	29	report (1)	报告	1	
19	30	engineering	工程	2	4
20	1	product	产品	1	
20	2	technology	技术	1	
20	3	information	信息	1	
20	4	trademark	商标	1	

Topic	Weight Rank	English	Word	# Topics	Overlap (other topics defined by same word)
20	5	register (1)	注册	2	6
20	6	use	使用	1	
20	7	advertisement	广告	1	
20	8	knowledge	知识	2	12
20	9	website	网站	2	6
20	10	marketing	销售	1	
20	11	network	网络	2	6
20	12	consumer	消费者	1	
20	13	industry and commerce	工商	1	
20	14	property ownership	产权	1	
20	15	quality	质量	1	
20	16	market	市场	3	15, 17
20	17	provide	提供	4	14, 15, 16
20	18	system (2)	系统	1	
20	19	enterprise	企业	4	10, 13, 19
20	20	tort	侵权	2	9
20	21	service	服务	2	15
20	22	protect	保护	2	8
20	23	content	内容	1	
20	24	problem	问题	14	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 16
20	25	carry out	进行	5	7, 10, 16, 19
20	26	nation	国家	6	3, 5, 7, 12, 19
20	27	time	时间	6	1, 2, 12, 13, 16
20	28	international	国际	2	5
20	29	operate	经营	3	10, 19
20	30	documentation	资料	1	