

データとしての習近平重要講話

2024年4月

ISS Discussion Paper Series

J-249

伊藤亜聖*, 林載桓†, 張紅詠‡

※本稿は伊藤亜聖「習近平政権の政策語彙～2012年11月15日-2022年10月23日～」日本国際問題研究所・中国研究会報告書, 27-56頁, 2023年3月刊行をもとに加筆修正したものである。本研究はJSPS科研費24K03158の助成を受けた。

* 東京大学社会科学研究所准教授: asei@iss.u-tokyo.ac.jp.

† 青山学院大学国際政治経済学部教授: t13025@aoyamagakuin.jp.

‡ 経済産業研究所上席研究員: zhang-hong-yong@rieti.go.jp.

要旨

中国の政治経済体制を理解する上で、最高権力者である習近平の影響力を実証的に分析することが必要とされている。本稿はウェブ上に公開されている習近平系列重要講話データベースを取り上げ、そのデータ上の特徴と利用上の注意点、そして分析事例を紹介する。同データベースは主に人民日報と新華社通信の記事によって構成されており、我々のリバーズエンジニアリングから習近平個人の日々の活動が収集されていることが確認された。キーワード検索の結果を見ると、同データと人民日報の間には相関関係があるものの、習近平系列重要講話データベースは習近平個人の意向をより早い時期に反映し得る。近年では同データベースを用いた実証研究が現れている。データベースには内容、サイズ、期間の面で限界がある一方で、習近平政権の実証研究に資する多頻度かつ高次元なデータとしての利用価値がある。

1. イントロダクション

本稿は2つの学術的潮流を受けて、中国共産党総書記の習近平に関わるテキストデータに着目し、その特性を検討する。第一の潮流は、近年の政治学および経済学における権威主義体制への関心の高まりである(Geddes, Wright, and Frantz, 2018; Egorov and Sonin, 2023)。グローバルに見て、2000年代以降に権威主義的な政治体制がむしろ増加する傾向が指摘されてきた。そして第二の学術的潮流は、コンピューターおよび機械学習の発達を背景とした量的テキスト分析の発展と普及である(Gentzkow, Kelly, and Taddy, 2019; Grimmer, Roberts, and Stewart, 2022)。権威主義体制の統治メカニズムを解明するうえでも、検閲に関する実証研究や(King, Pan, and Roberts, 2013)、政治リーダーのイデオロギーの変容の測定(Guriev and Treisman, 2019)といった論点で、テキストデータが新鮮な学術的な貢献をしてきた。しかし独裁者を含む政治家個人の影響力に関しては、データ上の限界からいまだに明らかになっていない論点が多い(Sinkkonen 2021)。

本稿は権威主義体制のなかでも、中国共産党の統治のもとで高度経済成長を実現してきた中国の事例、とりわけ習近平政権に注目する。2012年11月15日に中国共産党総書記と党中央軍事委員会主席に選出された習近平は大胆に政策決定過程全般に個人的な影響力を強めていった(Heilmann and Stefan 2016; Lampton 2015)。習近平は国家主席、共産党総書記、中央軍委員会主席(党および政府)に加えて、国家安全委員会主席や中央全面深化改革領導小組

組長を含む新旧の、そして多数の政策決定機構を率いてきた¹。それゆえにメディアや一部の研究者は習近平を「すべての主席」(the Chairman of everything)とも呼んだ(The Economist 2016; Hernández 2017; Shirk 2018; Economy 2018)。2010年代以降の中国の政治経済体制を理解する上で、最高権力者である習近平が及ぼす影響力への関心が格段に高まっている(Pei, 2022; Shirk, 2023)。

しかし共産党体制については、1980年代以来、政治家個人よりも組織や制度の変遷や影響を重視する観点が存在してきた。具体的には、政策形成においては、共産党と中央省庁に加え、国有企業や関連組織を含む多数のアクターからなる官僚組織の役割が強調されてきた(Lieberthal and Oksenberg, 1988; Jakobson and Knox, 2010)。そして政策実施の面では中央の「緩い統制」の下、地方政府が幅広い裁量を持つことが指摘されてきた(Ang, 2016)。習近平政権期の官僚組織や地方政府に関する分析においても、アクター間の利益相反と競争の側面を重視し、引き続き特定の官僚組織や地方政府が政策の形成と実施の両面で重要な役割を果たしているとする立場もある(Jones and Hameiri, 2023; Jaros and Pan, 2017; Jaros and Tan, 2020)。

それでは、「すべての主席」たる習近平はどの程度、政策決定過程や実体経済に影響力をもっているのだろうか。習近平の影響力を実証的かつ直接的に計測しようとする際に分析者が直面する問題は、習近平に関する多頻度かつアクセス可能な情報源が限られることである。この問題は中国に限られたものではない。権威主義体制において、最高権力者がいかに権限を行使し、政策過程や実体経済にいかなる影響を与えているかを考えるうえで、習近平に関する実証分析には中国研究を越えた意義がある。

本稿はこうした問題意識から、中国共産党新聞網のウェブ上に公開されている習近平系列重要講話データベース (*Xi Jinping Xilie Zhongyao Jianghua Shujuku*)に注目する(以下、Xi-Databaseと呼ぶ)²。同データベースは2012年11月15日以降の総書記在任中の習近平による講話、活動報告、現地視察、会見、外遊、祝電などの各種報道を合計1万件以上収録している。このテキストデータベースの利用例としては少数の記事を抽出した質的および量的な分析があるほか(Jiang, 2022; 邵, 2019)、より多くのサンプルを用いたフルテキストの量的分析例もある(伊藤, 2023; Ito, Lim, and Zhang, 2023a; Ito, Lim, and Zhang, 2023b; Lim, Ito, and Zhang, 2023; 邵, 2023)。

Xi-Databaseに着目することは、中国研究の伝統と近年の社会科学の発達の両面から見て、理に適うことである。テキストは共産圏研究にとって古典的な資料であった(Leites, Bernaut, and Garthoff, 1951; Oksenberg, 1969)。1980年代以降にも中国共産党の機関紙『人民日報』

¹ 習近平が主任・組長を務める政策決定機構はこの他に主要なものだけで中央外事工作領導小組組長、中央財經領導小組組長(領導小組はのちに委員会に改組)、中央网络安全和信息化委員会、中央全面依法治国委員会、中央軍民融合發展委員会を挙げることができる。

² <http://jhsjk.people.cn/> [最終閲覧 2024年2月12日]。

(*People's Daily*)を筆頭に継続的にテキストは分析に利用されてきた(Carlson et al, 2010; Johnston, 2013)。冒頭に述べたテキストデータの利用の普及も踏まえれば、Xi-Database を分析に利用することは、ごく自然な選択である。

しかし Xi-Database には注意点もある。本稿の課題はこうした注意点に着目し、Xi-Database を用いて有意義な分析を行ううえで重要な情報を提示することである。そもそも中国研究において、単一の情報源に依存することの限界は古くから指摘されてきた(Oksenberg, 1969)。加えて共産党体制下で「党の舌」としてプロパガンダの役割を担う新聞にバイアスが含まれる問題も、過去から現在まで度々指摘されてきた(岡部, 1966; Piotroski, Wong, and Zhang, 2017)。Xi-Database は習近平個人に関する報道を集めたものであるため、一般的な新聞とトピックの構成が異なるのは当然である。実務上、特に重要だと考えられるのは、Xi-Database に含まれる記事の元データがどのような特性のものかを理解し、データベースに含まれる重複記事の問題やその他の特性を理解することである。データのサイズや期間の短さといった限界を踏まえたうえで、我々は習近平政権を実証的に検討するうえで、Xi-Database に十分な価値があると主張する。

本稿の課題と構成は次のとおりである。まず第二節で Xi-Database の成り立ちと基本的な特徴を明らかにする。そこでは Xi-Database のデータソースの構成を確認する。特に、『人民日報』記事のうちでどのような記事が Xi-Database に掲載される傾向があるかを実証的に明らかにする。次に第三節で実際の分析利用例として辞書方式による分析を行う。国名と政策キーワードの頻度を確認したうえで、特に『人民日報』との比較においてどのような特徴が現れるかを検討する。第四節では、更なる分析の事例を紹介し、同時にデータベースの限界とその克服の可能性について議論する。最後に第五節で議論をまとめる。

2. Xi-Database の基本情報

2-1. データベースの成り立ち

Xi-Database に掲載されているのは、2012年11月15日、すなわち習近平が中国共産党総書記に就任して以降の『人民日報』(国内版、海外版)、『新華社通信』(*Xinhua News Agency*)、『求是』(*Qiushi*)、その他の同氏の関連記事である。具体的には同氏の講話、活動報告、現地視察、会見、外遊、祝電や弔電、その他の関連活動(例えば関連書籍の出版等)を報道した記事が収録されている。これら全記事はウェブ上に公開されている。Xi-Database の特徴として習近平個人の講話と行動に関する報道に重点がありながらも、同氏が出席した中国共産党と政府の重要会議の記事も含まれていることである。このため Xi-Database は習近平個人の動向に重点がありながらも、広い意味での党中央および中央政府の方針までが含まれるデータと捉えられる。

基本的な事実として、Xi-Database は人民日報社と中国共産党新聞網が開発し、2017年9月8日にリリースした。「習近平の一連の重要演説の精神をいつでもどこでも学べる」と述

べるリリース報道はデータベースを下記のように説明する³。

「習近平重要演説データベースには、習近平総書記が第18回全国代表大会以降に行った300以上の演説原文が収録されている。そして経済、政治、文化、社会、エコロジー、党建設、国防、外交など、さまざまな分野をカバーする記事は合計で6,000本を超える。」

これらのことから党中央の了承のもとで、人民日報社がデータベースを構築したことがわかる。外国語での記事が掲載されていないことなどから、Xi-Databaseの主要な目的は中国国内での宣伝であると考えられる。実際に地方政府が地元での政治的教育のためにXi-Databaseを利用すべきとの通知も見られる⁴。中国国内で、中国共産党高官に関する報道をまとめる公式ページが開設されることは珍しいことではない⁵。しかしここまで目立つ形で特定個人の記事集を示す例は他には見当たらない。

2-2. データベースの基本統計

本項ではXi-Databaseの基本的な構成を確認する。以下では習近平政権の第一期と第二期の期間、すなわち2012年11月15日から2022年10月23日を対象にデータを確認する。

表1にはダウンロード時点での12,511件の概要を示している。おおむね毎年1,000件程度の記事で構成されている。毎年の総記事数には若干の波があり、2014年が最大、そしてその次が2019年となり、2020年、2021年には記事数が減少している。また1記事あたりの平均文字数は同期平均で1,468字である。中央値は824字となっており、平均1,468字よりも短い記事が多いことを意味している。記事全体の78.1%に当たる9,776記事が2,000字以下である。逆に1万字を越える長文記事も少数ながら含まれており、全体の0.75%に当たる94

³ 「“习近平系列重要讲话数据库”正式上线」人民网-中国共产党新闻网 2017年9月8日記事(<http://dangjian.people.com.cn/n1/2017/0908/c117092-29524442.html>)、[最終閲覧 2024年2月7日]。

⁴ 「关于用好“习近平系列重要讲话数据库”开展学习教育的通知」、中共许昌市教育局直属机关委员会、2017年9月12日(<http://jyj.xuchang.gov.cn/lxyz//8d72f5d7-3ac8-4bd4-bff6-e47ef53774d9.html>) [最終閲覧 2024年2月7日]。

⁵ 中国共産党新聞網には高官に関する報道をまとめたページがある。例えば李強(Li Qiang) (<http://cpc.people.com.cn/GB/64192/448521/index.html>)、蔡奇(Cai Qi) (<http://cpc.people.com.cn/GB/64192/448522/index.html>)のページを参照。また国務院総理である李強に関しては国務院新聞弁公室サイトにも特設サイトがある (http://www.scio.gov.cn/yw/lq/_index_3.html) [3記事とも最終閲覧 2024年2月7日]。同種の情報は李克強時代にもあったが、首相退任に伴ってアクセスできなくなっているものが少なくない。

件が該当する。平均文字数に関しては2019年以降、若干短くなっている⁶。

補足表1に示したように、記事の情報源は合計47種類ある。このなかで『人民日報』および『人民網』(People's Daily Online)はそのうち6,764件を、『新華社通信』および『新華網』(Xinhua Net)は5,438件を占めている。つまり、特筆に値する事実として、ウェブ版を含めた二大国有メディア企業の記事によってXi-Database全体の97.5%が構成されている。これに続くのは中国共産党の機関雑誌『求是』であるが、その数は80件にとどまり、政権第二期に入って以降に記事数が増加している。

表1 Xi-Databaseの概要

年	記事数					総文字数	1記事の 平均文字数
	小計	人民日報 &人民網	新華社通信 &新華網	求是	その他		
2012	66	35	31	0	0	111,078	1,683
2013	1,148	549	559	1	39	1,440,400	1,255
2014	1,652	848	717	0	87	2,527,592	1,530
2015	1,381	792	560	1	28	2,144,458	1,553
2016	1,184	725	440	0	19	2,101,857	1,775
2017	1,094	649	428	1	16	1,901,084	1,738
2018	1,297	782	507	1	7	2,001,699	1,543
2019	1,542	768	748	13	13	1,943,128	1,260
2020	1,320	639	653	26	2	1,832,732	1,388
2021	1,034	552	457	18	7	1,383,127	1,338
2022	793	425	338	19	11	975,210	1,230
Total	12,511	6,764	5,438	80	229	18,362,365	1,468

注：集計対象は2012年11月15日-2022年10月23日。クリーニング前のデータである。

『人民日報』と『人民網』の6,764件のうち、571件は『人民網』(<http://www.people.com.cn/>)関連の記事で、残りの6,193件が海外版を含む『人民日報』である。

出所：Xi-Databaseより筆者ら作成。

⁶ 2012年12月のいわゆる「中央の八項目規定」通知に、中央政治局員に関する報道をより短縮するとの記載があったが、仮にこの通知が効果を持っていたのであれば2019年以降に平均記事数が減るのは不自然である。2012年12月05日、中国共産党新聞網「中共中央政治局召开会议 习近平主持」(<http://cpc.people.com.cn/n/2012/1205/c64094-19793530.html>)を参照[最終閲覧2024年2月7日]。

2-3. リバースエンジニアリング

Xi-Database をデータとして扱ううえで検討すべき論点の一つは、記事の収集メカニズムである。『人民日報』と『新華社通信』に掲載される習近平に直接かかわる記事が収集されていると考えられる一方で、このプロセスを完全に明らかにすることは難しい。特に『新華社通信』の場合、第三者が利用できる同社記事データベースは存在せず、どのような記事が Xi-Database に収集されているかを実証的に検討することは困難である。それに対して『人民日報』の場合には、フルテキストデータを提供するデータベースが存在する。

以下では『人民日報』のフルテキストデータを用いて、Xi-Database にどのような記事が掲載されやすいかを実証的に検証する。『人民日報』のデータは2012年11月15日から2022年8月15日までで、合計320,925件である⁷。同期間の Xi-Database に掲載されている『人民日報』関連の記事は6,098件である。つまり『人民日報』掲載記事の1.9%(6,098/320,925)が Xi-Database に掲載されている計算になる。

我々はこれらの情報を用いて Xi-Database の生成過程をリバースエンジニアリングする⁸。我々は両データベースの記事タイトルをもとに、Xi-Database 掲載の6,098件のうち、85%に当たる5,180件について『人民日報』データとマッチングすることが出来た。これにより、『人民日報』データの320,925件のうち、Xi-Database に収集されている5,180件については1を、それ以外の記事では0となる被説明変数を作成できる。ここでフルテキストデータを用いて、どのような単語が含まれている記事が Xi-database に含まれやすいかをテキスト回帰(text regression)によって推計することができる⁹。テキスト回帰はその名の通り、テキストに含まれるデータ、この場合には単語を説明変数とした回帰分析である(Gentzkow, Kelly, and Taddy, 2019)。被説明変数に影響を与える効果が薄い変数(単語)については、Lasso 推計により係数を0とする penalty をかけることで、重要な変数(単語)のみを抽出することができる¹⁰。

⁷ 『人民日報』データは Oripote Information Services, Inc. (<https://www.oripote.com/>)より得た。

⁸ テキスト回帰のほかのアプローチとして、トピックモデルを推計し、どのようなトピックの記事が収集されやすいかを検討することも考えられる。本稿ではよりミクロな規定要因に着目するためテキスト回帰を採用する。

⁹ 推計の際の統計処理は R、特に `quanteda package` と `glmnet package` によって実施した (Benoit, et al., 2018; Friedman, et al., 2023)。語彙を分かち書き(segmentation)する必要があるが、我々はこのプロセスで `jiebaR` を用いた。

¹⁰ Penalty の大きさを調整する Lambda (λ)ごとに推計ができるが、ここではとくに少数の単語を抽出するために、比較的大きな Lambda を設定して、重要な変数を選別した。なお、モデル全体を最適化する際にはデータスプリットを行い、交差検証をする等の手続きにより最適な Lambda を特定することが望ましい。ただし、本稿での課題は少数の主要な語彙を抽

表2にテキスト回帰の結果を示した。パネルAには、記事のタイトルをデータとして分析した結果を、そしてパネルBには記事の本文をデータとして分析した結果を示している。まずパネルAの結果から、記事タイトルに習近平や常務会議、そして中国共産党中央政治局が含まれている記事が高い確率で掲載されていることがわかる。パネルBから、第一に、『人民日報』と『新華社通信』の特定の記者名(杜尚泽, 兰红光)や専属カメラマンの氏名(李学仁, 谢环驰, 鞠鹏, 庞兴雷)が含まれる記事がXi-Databaseに掲載される傾向が高いことがわかる。第二に常務会議、中央軍事委員会といった重要会議、そして第三に釣魚台迎賓館、外交儀礼、祝電、慰問電信といった外交活動に関わる用語が含まれている。そして第四に重要な政治イベントに同行していると考えられる特定の高位政治家(王沪宁, (丁)薛祥, 杨洁篪)が登場している。

これらの結果から、Xi-Databaseには、『人民日報』記事のうちで、特に習近平個人の日々の政治的活動を収集していることが確認できる。これらの結果は驚くに値するものではないが、『人民日報』と『新華社通信』の特定記者名や専属カメラマンに関わる、ある種の特別にクラス分けされた記事が収集されていると考えることができる。

表2 Lasso 推計結果による単語レベル係数

パネル (A) タイトル				パネル (B) 本文			
Lambda=0.02		Lambda=0.005		Lambda=0.02		Lambda=0.008	
単語	係数	単語	係数	単語	係数	単語	係数
习近平	0.417	常务会议	0.822	杜尚泽	0.240	杜尚泽	0.306
常务会议	0.384	习近平	0.622	王沪宁	0.114	晚同	0.245
		主席令	0.545	薛祥	0.039	王沪宁	0.140
		中共中央 政治局	0.333	诚挚	0.029	另发	0.135
		重要讲话	-0.235	愿同	0.023	惊悉	0.122
		隆重举行	0.232	新华社	0.017	李学仁	0.120
		茶话会	0.230	北京人民 大会堂	0.011	薛祥	0.119
		指引	-0.192	杨洁篪	0.003	谢环驰	0.103
		特使	-0.182	习近平	0.001	我同	0.099
		军衔	0.160			鞠鹏	0.092
		党外人士	0.087			诚挚	0.078

出ることであるので、これらのモデル最適化の手続きを実施していない。

		贺信	0.082			北京人民 大会堂	0.074
		学习	-0.070			三要	0.053
		思想	-0.070			日同	0.044
		致贺	0.070			愿同	0.039
		团拜会	0.067			我愿同	0.032
		深入	-0.067			特命	0.032
		社会主义	-0.059			兰红光	0.028
		贯彻	-0.049			贺电	0.028
		特色	-0.046			杨洁篪	0.025
		总统	0.036			新华社	0.024
		就职仪式	-0.033			常务会议	0.018
		论述	-0.027			慰问电	0.015
		发表	0.021			祝贺	0.015
		国事访问	0.019			致以	0.012
		通电话	0.011			庞兴雷	0.009
		届	0.009			北京钓鱼 台国宾馆	0.006
		会见	0.009			欢迎仪式	0.005
		系列	-0.004			两国人民	0.005
		驻外	0.002			中央军委	0.004
		时代	-0.002			习近平	0.002
						出席	0.001
						全权大使	0.001

出所：Xi-Database より筆者ら推計。

2-4. クリーニング上の論点

Xi-Database を分析する際には、一般的なテキストデータと同様にクリーニングが必要である。一つ本データに特徴的なのは、同一記事が『人民日報』と『新華社通信』に掲載されることが頻繁にある点である。我々の分析によれば、全文が同一の記事が69件あり、さらに前後14日間でタイトルが同一のものが2,314件ある。これらを除くとデータ総数は10,197件となる。『人民日報』と『新華社通信』では、若干のタイトルの変更をしている場合もあるため、より厳密に重複をカットする場合には、文書間の類似性が極端に高い記事をカットするという手段も考えられる。ただし、そもそも重複記事を落とすべきか否かは分析課題次第である。なぜなら『人民日報』と『新華社通信』に同時に掲載される記事はそれだけ政権

にとって重要な情報であるとも考えられるためである¹¹。

分析に当たっては課題次第で、さらにクリーニングを行うことが望ましい。例えば極端に短い記事は、例えば写真に一行の説明がついているだけのものもあり、分析上の意味は薄いと考えられる¹²。また逆に極端に長い記事は、共産党大会における長文の演説などが含まれる。このため分析課題によってはカットすることもあり得る¹³。この他に例えば習近平の演説のみに着目する場合には当該記事を抽出することも考えられる。また『求是』記事は一般的な新聞記事と性質が異なる面もあるため、分析課題によってはカットすることも考えられる。

3. キーワード検索

3-1. 分析上の想定と論点

テキストデータから政策キーワードや何らかの尺度を抽出する意図は、言及頻度からある程度、政策の実施状況が推測できたり、あるいは、何らかの効果が表れると想定するためである。しかしここには仮定があり、それは発言の有無と実行(あるいは効果)の有無との間で、有言実行の関係があると想定していることになる。発言の有無と実行の有無の間には、表3に示した4つの可能性があり、有言実行を想定した場合には、発言の有無や頻度から、政策実施の状況を推定することができる。発言も無く、実行も無い場合、すなわち表3の不言無行は分析対象外となるため問題が生じない。しかし政策実施とのつながりを想定した際に問題となるのは、「発言があるのに、政策実行を伴わない」という有口無行と、「発言はしないにも関わらず、政策が実行されている」という不言実行の場合である。有口無行と不言実行を含めた検討は第4節で行うこととして、ひとまず第3節ではテキストデータからのキーワードの抽出結果を報告する。

¹¹ 例えば外交部部長と国務院副総理を務めた銭其琛(Qian Qichen)の葬儀に関する記事は『人民日報』と『新華社通信』の両方に掲載されている。2017年5月19日『人民日報』出所記事「钱其琛同志遗体在京火化 习近平等到八宝山革命公墓送别」(<http://jhsjk.people.cn/article/29285636>)および2017年5月18日『新華社通信』出所記事「钱其琛同志遗体在京火化 习近平等到八宝山革命公墓送别」(<http://jhsjk.people.cn/article/29285117>) [両記事とも最終閲覧2024年2月12日]。

¹² 例えば100文字未満の記事数は425件であり、全体の3.4%である。ごく短い記事の事例としては2016年7月26日記事「组图：七月习近平精彩镜头全纪录」(<http://jhsjk.people.cn/article/28586429>)や2020年4月21日記事「习近平点赞柞水木耳」(<http://jhsjk.people.cn/article/31681383>)がある [両記事とも最終閲覧2024年2月12日]。

¹³ 最も長文なものは19回党大会における習近平による著名な基調報告である。2017年10月28日記事「习近平在中国共产党第十九次全国代表大会上的报告」(<http://jhsjk.people.cn/article/29613660>) [最終閲覧2024年2月12日]。

表3 発言と実行・効果のマトリックス

		実行または効果の有無	
		有	無
発言の有無	有	有言実行 (発言有り、効果あり)	有口無行 (発言有り、効果なし)
	無	不言実行 (発言なし、効果あり)	不言無行 (発言なし、効果なし)

出所：筆者ら作成。

3-2. 国と地域

以下では Xi-Database から基本的な情報を抽出して検討する試みとして、①国名・地域名の抽出、②政策名の抽出、③『人民日報』記事との比較を行う。分析に用いるデータは、重複記事の削除(全文一致記事のカットおよび前後 14 日間での同一タイトルカット)、極端に短い記事と長い記事をカット(100 字未満、1 万字以上の記事カット)した 9,725 記事である。

まずテキストデータに登場する国・地域別の名称(香港、マカオ、台湾を含む)をカウントした¹⁴。表 4 には通期の上位 15 の国・地域について、各年の言及回数を示した。まず 1 年ごとの国・地域の言及回数の推移を見ると、2014 年に合計 2,218 回となっており、最多を記録し、その後 2017-2018 年には 1,000 回程度となっていたが、2020 年および 2021 年には 600 回へと大きく減少している。国・地域への総言及回数(A)を年ごとの総記事数(B)で割ると、1 記事当たりで平均していくつの国・地域に言及しているかを測れる。この値で見ても 2014 年の 1.69、2018 年の 1.25 か国(地域)から、2020 年には 0.59 か国(地域)へと言及頻度が大きく低下している。

¹⁴ 国名・地域名のカウントの際には、R のパッケージ Newsmap(Watanabe, 2018)の 241 か国の簡体字の国名リスト

(https://github.com/koheiw/newsmap/blob/master/dict/chinese_simplified.yml)をもとに、国名を再検討した。中国政府の公式な言及方法を確認する際には中国外交部 HP を参照した。Newsmap は国名・都市名から特定記事がどの国について言及している蓋然性が高いかを推定するパッケージであるが、本分析のタスクにおいて精度上の課題があったため、直接国名・地域名を辞書方式で抽出する方法を採用した。

表 4. 国・地域別の登場記事数(上位 15 カ国・地域)

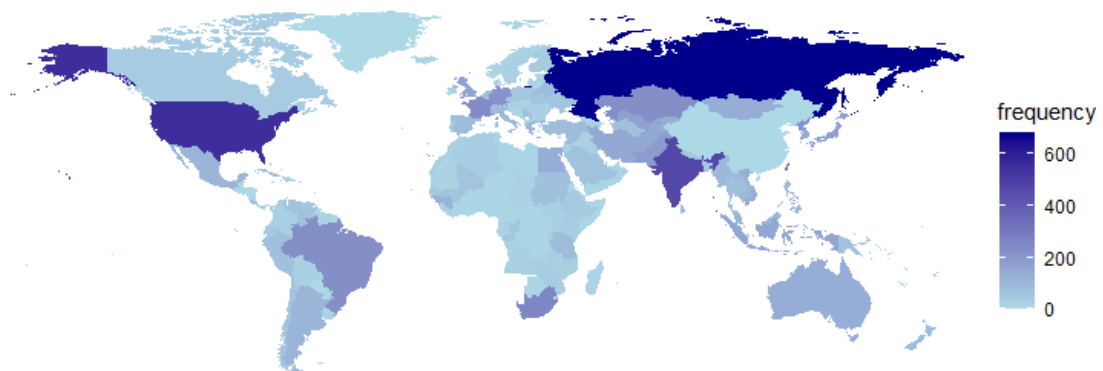
国名	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	総計
ロシア	2	123	110	101	52	53	68	82	30	37	24	682
米国	4	97	104	91	69	65	37	28	26	21	13	555
インド	0	79	113	62	49	28	49	49	16	11	13	469
香港	2	32	40	32	18	63	42	42	49	37	43	400
マカオ	3	24	59	25	17	41	40	62	29	28	27	355
台湾	2	35	44	37	30	26	38	36	23	25	23	319
南アフリカ	0	60	26	54	13	22	56	10	8	4	7	260
フランス	0	9	62	33	16	12	20	51	22	13	5	243
ブラジル	0	31	75	34	17	17	12	29	4	5	7	231
カザフスタン	0	50	30	40	12	26	16	24	2	7	24	231
北朝鮮	0	23	33	20	21	34	29	37	13	12	2	224
ドイツ	1	13	56	26	15	44	12	20	18	11	4	220
パキスタン	0	14	27	45	21	11	18	20	10	12	9	187
日本	0	12	34	42	8	13	12	32	13	7	7	180
ウズベキスタン	0	40	19	13	27	6	9	18	2	7	30	171
241 国・地域への 総言及回数(A)	18	1,553	2,218	1,505	1,095	971	1,249	1,330	593	649	546	11,727
年総記事数(B)	59	943	1,310	1,082	922	847	999	1,149	1,007	801	606	9,725
(A)/(B)	0.31	1.65	1.69	1.39	1.19	1.15	1.25	1.16	0.59	0.81	0.90	1.21

注：集計期間は 2012 年 11 月 15 日-2022 年 10 月 23 日。

出所：Xi-Database より筆者ら作成。

国別の言及頻度を見ると、通期で見て、最多の言及回数は上位からロシア、米国、インド、香港、マカオとなっている。通期で見ると、BRICS 諸国との関係が比較的に目立つ。ただし一定の変動も見られる。2016 年には米国が最多の言及回数となっていた。しかし米国との外交関係が悪化した 2018、2019 年にはロシアが再び最多となっている。2013 年から 2015 年のもう一つの特徴は、一帯一路構想と関連する中央アジア諸国の回数が相対的に増えたことである。例えば 2015 年には、期間総合の順位で 10 位のカザフスタン、13 位のパキスタンと言った国々への言及が相対的に多かった。2020 年以降を見ると、総回数の減少の一方で、香港への言及回数が高止まりしており、2020 年、2021 年、2022 年では香港が最多の回数となっている。図 1 は通期の言及回数を地図上に示した。

図 1. 国・地域別の登場記事数



注：中国は値を 0 としてある。集計期間は 2012 年 11 月 15 日-2022 年 10 月 23 日。

出所：Xi-Database より筆者ら作成。

3-3. 政策キーワード

続いて経済、内政、外交、安全保障の四分野について、補足表 2 に示した 71 のキーワードの登場する記事数を集計し、補足表 3 にその結果を示した。まず補足表 3 の (B) 列に期間通算での該当記事比率(全期間での当該キーワードの登場記事数/全期間での全記事数)を示して、その順位で並び替えている。この(B)列は分析期間を通じた言及頻度を意味しており、予想されるように、「偉大な復興」、「中国の夢」といったスローガンが 1 位と 3 位となっている。「運命共同体」、「一帯一路」、「小康社会」も上位に入っている。また表では第一期と第二期での言及比率も示している。上位のキーワードのなかでは、「偉大な復興」は期間を通じて言及頻度が大きく変化していないのに対して、「中国の夢」への言及頻度は第二期には低下している。「一帯一路」、「運命共同体」は第二期にとくに言及頻度が高まっている。

次に単月での最大記事比率となった月(補足表 3 の(E)列)を基に、ピークを迎えた月ごと

にキーワードをカレンダーに落とし込んだものが表5である。単月での最大記事比率は、つまりデータをカバーする10年間で最も当該キーワードが言及された月を意味するため、「瞬間最大風速」に当たる。総じて任期期間中に、各キーワードはばらばらに各々のピークを迎えていることが確認でき、キーワードの流行りと廃りが見て取れる。「瞬間最大風速」は、当然ながら政治・外交的なイベントと深く関係があることは明らかだろう。例えば12月に経済系のキーワード(マクロ調整、不動産、城鎮化、公平競争、反独占)が目立つのは、毎年12月に中央経済工作会議が開催され、関連記事が増加するためだと考えられる。同様に、3月には两会(全国人民代表大会および中国人民政治協商会議)が開催されるため、2-3月は两会に向けたキーワードが対応する傾向がある。

表 5. 習近平政権のキーワードカレンダー(月次登場比率が最大の月に表示)

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
2012	-	-	-	-	-	-	-	-	-	-	小康社会, 戦略機遇期	マクロ調整, 不動産
2013			韜光養晦									城鎮化, 糧食安全, 決定的作用
2014			混合所有制		网络安全						AIIB	
2015	反腐	法治	新常态	核心的利益				純潔	国際連合, 新型大国関係	世界銀行, 中国製造 2025	周辺外交, IMF	
2016				就業問題				高齢化	G20			二つの百年
2017	保護主義	新疆, 経済安全	供給側改革						BRICS	偉大な復興, 総体国家安全観		ネットワーク主権
2018	大衆創業, 収入格差	中国夢	軍民融合		断絶的技術	SCO, 新型国際関係	運命共同体				WTO	
2019		憲政		一帯一路								いじめ主義
2020		減税		戦略的新興産業				デジタル化			一国主義, 双循環, RCEP, データ安全	公平競争, 反独占
2021		脱貧困, 台湾, 新発展段階			基礎研究, 大有作為, 新型拳国体制		炭素中立			平台独占	TPP	
2022				リスク解消		エネルギー安全	香港, 動態清零	共同富裕		国家安全	-	-

注: 集計期間は 2012 年 11 月 15 日-2022 年 10 月 23 日。

出所: Xi-Database より筆者ら作成。

3-4. 『人民日報』との比較

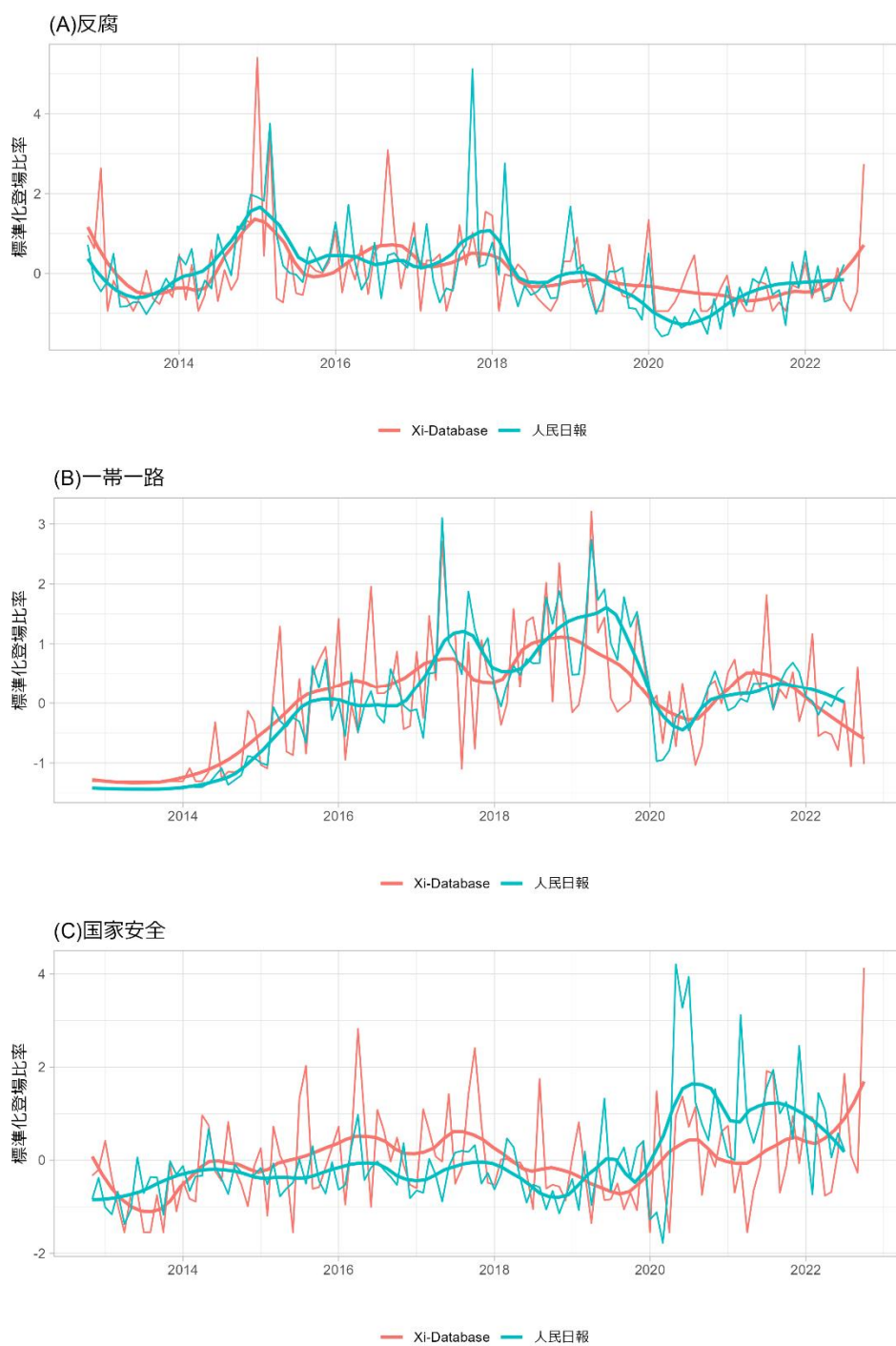
次にいくつかのキーワードを用いて『人民日報』と Xi-Database の該当記事数を比較する。図 2 には習近平個人の意向が比較的強く反映されていると考えられる政策キーワードとして、「反腐」、「一带一路」、「国家安全」を取り上げた。図の縦軸は標準化された月次該当記事比率である。例えばある月の『人民日報』の記事総数に占める「反腐」を含む記事数の比率を求め、それを標準化(平均 0、分散 1)したものである。

ここから判明するのは、第一に、いずれのキーワードに関しても両データの間に関連は比較的高いことである。Xi-Database に含まれている記事の約半数が『人民日報』から提供されていることから、これはある意味で当然のことである。しかしキーワードによって、関連の強さには差がある点は注目に値する。「一带一路」では相関係数 0.79 と強く同期しているのに対して、「反腐」では 0.54、「国家安全」では 0.30 とより弱い相関となっている。ここで Xi-Database が習近平個人の言動と行動を反映するのに対して、『人民日報』が習近平個人ではなく中国共産党中央の官僚機構としての政策優先順位を反映していると仮定してみよう。その場合には、「一带一路」は習近平個人と官僚機構の政策優先順位がかなり一致していると考えられる一方で、「国家安全」を巡っては両者の間に一定の意見的差異が見られた可能性がある。

第二に注目に値することは、Xi-Database がより早い時期に強い振幅を見せている点である。これは特に「一带一路」と「国家安全」で特徴的である。「一带一路」の場合、2015 年から 2016 年にかけて Xi-Database が強く反応しており、この時期はまさに「一带一路」構想を政権として目玉の外交政策とするために注力していた時期に当たる。「国家安全」に関して興味深いのは、Xi-Database において 2015 年から 2017 年にかけて強く反応している月が少なくない点である。『人民日報』では 2020 年以降になってようやく強い反応が見られるが、習近平はより早い時期から「国家安全」に言及していた。

以上の我々の整理が示唆するのは、Xi-Database が『人民日報』と相関関係を持ちながらも、一部の政策論点については、Xi-Database がより早い時期に習近平個人の政策的優先順位に反応し、より早くシグナルを発信している可能性である。

図2. 辞書方式による Xi-database と『人民日報』の比較



注: 月次データの『人民日報』と Xi-Database の相関は、「反腐」で 0.547、「一带一路」で 0.797、「国家安全」で 0.301 である。

出所: 筆者ら作成。

4. 応用例と展望

4-1. 応用例

以下では Xi-Database を用いたより踏み込んだ分析を加えている論文を紹介する。

第一は、Xi-Database から習近平の政治的思想や政策アジェンダを抽出しようとする取り組みである。

一つの事例は邵(2023)である。著者は深層学習の一種(CNN)を用いて、Xi-Database の記事を①毛沢東主義、②ポスト毛沢東主義、③闘争的態度の3つの次元でそれぞれ分類をしている。著者はこの作業を通じて Xi-Database から習近平の思想を推計することを目指しており、その結果、時を追うごとに徐々に毛沢東思想に近づいていると結論付けている。また闘争的態度も強まっていると結論付けている。

より包括的に習近平政権の政策アジェンダを抽出しようとした事例として、構造トピックモデルで分析した Lim, Ito, and Zhang(2023)がある。トピックモデルはテキストデータ内に含まれる主要な話題を抽出する教師なし学習の手法である。彼らは Xi-Database のデータ内に、25 程度の政策領域が含まれていることを報告しており、大きく分けて内政トピックス、外政トピックス、そして危機対応トピックスによって構成されている。同論文では各トピックが 2012 年から 2022 年までどのように推移を示しており、例えば反腐敗運動に関する言及のピークが 2015 年から 2017 年頃にあったこと、一带一路への言及のピークが 2016 年と 2019 年の二度あったことなどを報告している。一带一路のトピック比率の推移と、中国による対外援助額を比較すると、対外援助額は 2016 年に一つのみのピークを示している。これは習近平政権が一带一路への言及を 2019 年にも重点化したのに対して、金融面での実際の行動(この場合は援助額)は右肩下がり続けていたことを意味する。つまり発言と行動の間には大きなギャップが生まれうることを示している。

第二に、Xi-Database から不確実性や政策的な志向を推計し、それらが実体経済にどのように影響しているかを推計した研究がある。

その一例は Ito, Lim, and Zhang(2023a)である。著者らは Xi-Database から経済政策不確実性(Economic Policy Uncertainty, EPU)指数を抽出している。EPU は新聞記事をデータとして用いることで、既存の不確実性指数である Volatility Index(VIX)等よりも長期間かつより多くの国々について計測できる手法である(Baker, Bloom, and Davis, 2016)。計測手法としては辞書方式(dictionary method)が採用されており、幅広く応用が進んでいる。Ito, Lim, and Zhang(2023a)は Xi-Database から EPU を計測することで、習近平個人の認識を反映した EPU、すなわち XiEPU の構築を目指した。その結果によると、XiEPU は既存の EPU 指数と一定の相関関係を見せた。そのうえで、著者らは上場企業のミクロデータと統合し、他の不確実性指標をコントロール変数しても、XiEPU の高まりは企業の投資行動に負の影響があることを報告している。Xi-Database が企業の意思決定にとって重要な情報を含んでいることを示唆する結果と言える。

もう一つの事例は、Xi-Database と国务院総理のテキストデータを比較することで、主要

政治家間の政策的志向の差を比較した Ito, Lim, and Zhang(2023b)である。著者らが着目するのは習近平と 2013 年 3 月から 2023 年 3 月まで国務院総理を務めた李克強の間での政策対立である。同論文は国有企業と民間企業への政策志向の違いに着目し、半教師付き機械学習の一種である Latent Semantic Scaling によって国家—市場指向指数(State-Market Orientation Index)を計測している。その結果によると、習近平は李克強よりも総じて国家指向が強く、市場指向の李克強との間の思考の差は 2015 年から 2016 年にかけて最大となった。習近平のトーンが比較的安定的であるのに対して、李克強の国家—市場指向指数は大きく波を打っている。分析によれば李首相が重視した創業支援政策(Mass Entrepreneurship, Mass Innovation Policy)を大々的に推進する週(Mass Innovation Week)には大きく市場指向となっていることを報告している。彼らは習近平と李克強の間の政策志向の差を一種の政治的不確実性と捉え、政策志向の差が大きくなると特に国有企業の投資に負の影響があることを報告している。

4-2. データの限界と更なる可能性

このように Xi-Database を用いることで、習近平の政策アジェンダ、政策的志向、企業行動への影響といった論点を実証的に検証することが可能となる。同時に、Xi-Database には分析上の限界があることも否定できない。

第一は、内容上のメディアバイアスである。言うまでもなく Xi-Database はプロパガンダの目的として作成されているため、ここから抽出可能な指数にもこうしたバイアスが含まれてしまう。ただし、この問題は、重要データである『人民日報』においても同様に直面する問題である。また、同時に注目が必要なことは、すでに紹介した研究が示唆するように、こうしたメディアバイアスが含まれているにも関わらず、Xi-Database に含まれる記事は実体経済や企業行動に対して影響力を持っていると考えられることである。

第二はデータサイズの限定性である。Xi-Database は表 1 に示したように、10 年間で元データが 12,511 件である。ここから重複記事を削除すると 10,197 件となる。平均で 1 年あたり約 1000 記事となり、テキストデータとしては決して大規模なデータセットとは言えない。多頻度のデータ、例えば日次レベルや週次レベルの指数を作成するには向いていない。一方で、月次であればそれなりに意義のある指数を作成することができるし、また多くの機械学習を実施する上では最低限のデータサイズを満たしている。

第三はカバー期間の短さである。集計期間は習近平が中国共産党総書記に就任した 2012 年 11 月 15 日以降に限られる。『人民日報』データであれば、1946 年以降の長期にわたりデータを得ることができる。『人民日報』の記事スタイルが 1980 年代に大きく変化することを踏まえても、習近平政権以前の胡錦涛政権期や江沢民政権期までを含めた中長期の指数(例えば反腐敗に対する態度)を作成するには『人民日報』やその他の新聞はより有力なデータソースとなる。ただしこの点については、本稿のリバースエンジニアリングの箇所でも実施したような手続きをもとに、通時的に通用する一定のキーワードを特定することで、例えば

疑似的な『胡錦濤重要講話データベース』を構築し、Xi-Database を比較することも考えられる。

このような限定性があることを踏まえたうえで、我々は、政策決定者としての習近平が、政策決定過程や実体経済にいかなる影響を与えているかを実証的に検討するうえで、Xi-Database に十分な価値があると考え。とくに言説と政策実行の間の関係を実証的に分析する上では、政策関連文書との結合には潜在的な価値がある。この他にも前述した関連研究に見られるように上場企業データとの統合を始めとして、他のマイクロデータと統合することで、権力者の影響力をより多面的に検討することが考えられる。

5. 結論

本稿はウェブ上に公開されている Xi-Database を取り上げ、そのデータ上の特徴と利用上の注意点、そして分析事例を紹介した。Xi-Database はその大部分が『人民日報』と『新華社通信』の記事によって構成されており、第二節で実施した我々のリバースエンジニアリングからも習近平個人の活動に関する報道が収集されていることが確認された。また反腐敗運動、一帯一路、そして国家安全を辞書方式によって計測し、Xi-Database と『人民日報』を比較した。その結果、両者には相関関係があるものの、キーワードによってその相関の強さは差があり、また一部キーワードでは Xi-Database がより早い時期に強いシグナルを発信している可能性が示唆された。近年、Xi-Database を用いた実証研究が複数現れており、その中には習近平思想や政策アジェンダに着目するもの、経済政策不確実性や政策志向の計測とそれら指数が企業行動に与える影響に着目するものがある。Xi-Database は、その構築目的からしても、多分にプロパガンダが含まれ、またデータのサイズと期間において限界があることは明らかである。これらの限界があるものの、我々は Xi-Database が習近平政権の実証研究に資する多頻度かつ高次元なデータとしてユニークな利用価値があると考えている。

参考文献

- Ang, Y., (2016). *How China Escaped Poverty Trap*. Cornell University Press.
- Baker, S. R., Bloom, N., & Davis, S. J. (2016). “Measuring economic policy uncertainty,” *The Quarterly Journal of Economics*, 131(4), 1593-1636.
- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). “quanteda: An R package for the quantitative analysis of textual data,” *Journal of Open Source Software*, 3(30), 774-774.
- Carlson, A., Gallagher, M. E., Lieberthal, K., & Manion, M. (Eds.). (2010). *Contemporary Chinese Politics: New sources, methods, and field strategies*. Cambridge University Press.
- Economy, E. 2018. *The Third Revolution: Xi Jinping and the New Chinese State*. New York: Oxford University Press.
- Egorov, G., & Sonin, K. (2023). The political economics of non-democracy. *Journal of Economic*

- Perspective*, forthcoming.
- Friedman, J., Hastie, T., Tibshirani, R., Narasimhan, B., Tay, K., Simon, N., & Yang, J. (2023). “glmnet: Lasso and elastic-net regularized generalized linear models,” *Astrophysics Source Code Library*, ascl-2308.
- Geddes, B., Wright, J., & Frantz, E. (2018). *How Dictators Work: Power, Personalization, and Collapse*. Cambridge, UK: Cambridge University Press.
- Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3), 535-574.
- Grimmer, J., Roberts, M. E., & Stewart, B. M. (2022). *Text as data: A new framework for machine learning and the social sciences*. Princeton University Press.
- Guriey, S., & Treisman, D. (2019). Informational autocrats. *Journal of economic perspectives*, 33(4), 100-127.
- Heilmann, S., & Stepan, M. eds. (2016). "China's Core Executive: Leadership Styles, Structures, Processes under Xi Jinping," Mercator Institute for China Studies.
- Hernández, J. C. (2017) “China’s ‘Chairman of Everything’: Behind Xi Jinping’s Many Titles,” *The New York Times*, 10月25th, 2017.
- 伊藤亜聖 (Ito, A.) (2023)「習近平政權の政策語彙～2012年11月15日-2022年10月23日～」日本国際問題研究所・中国研究会報告書, 27-56頁
- Ito, A., Lim, J., & Zhang, H. (2023a). “Catching the Political Leader’s Signal: Economic policy uncertainty and firm investment in China,” *China Economic Review*, 102035.
- Ito, A., Lim, J., & Zhang, H. (2023b). “Remembering Li Keqiang: From divergence to convergence in Zhongnanhai,” Harvard-Yenching Institute Working Paper Series, December 20, 2023.
- Jakobson, L., & D. Knox. (2010). “New foreign policy actors in China,” Stockholm International Peace Research Institute, SIPRI Policy Paper No.26.
- Jaros, K., & Pan, J. (2017). “China’s Newsmakers: Official Media Coverage and Political Shifts in the Xi Jinping Era,” *The China Quarterly*, 233.
- Jaros, K. A., & Tan, Y. (2020). “Provincial power in a centralizing China: the politics of domestic and international “development space”,” *The China Journal*, 83(1), 79-104.
- Jones, L. & Hameiri, S. (2021). *Fractured China: How State Transformation is Shaping China’s Rise*. Oxford University Press.
- Jiang, Y. (2022). “Unpacking the Belt and Road Initiative: Does Its Public Diplomacy Narratives Match Its Implementation?,” *East Asia*, 39(4), 315-331.
- Johnston, A. I. (2013). “How new and assertive is China's new assertiveness?,” *International Security*, 37(4), 7-48.
- King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *American political science Review*, 107(2), 326-343.

- Lampton, D. M. (2015). "Xi Jinping and the National Security Commission: Policy Coordination and Political Power," *Journal of Contemporary China*, 24(95): 759-777.
- Leites, N., Bernaut, E., & Garthoff, R. L. (1951). "Politburo images of Stalin," *World Politics* 3.3: 317-339.
- Lieberthal, K. & Oksenberg, M. (1988). *Policy Making in China: Leaders, Structures, and Processes*. Princeton University Press.
- Lim, J., Ito, A., & Zhang, H. (2023). "Policy Agenda and Trajectory of the Xi Jinping Administration: Textual evidence from 2012 to 2022," RIETI Policy Discussion Paper Series, 23-P-008.
- 岡部達味(Okabe, T.) (1964). 「内容分析による中共対外政策の研究」『アジア研究』10(4)、28-58 頁。
- Oksenberg, M. (1969). "Sources and methodological problem in the study of contemporary China," in A. Doak Barnett ed. *Chinese Communist Politics in Action*, Seattle and London: University of Washington Press. pp.577-605.
- Pei, M. (2022). "Xi Jinping's Political Agenda and Leadership: What Do We Know from His Decade in Power," *China Leadership Monitor*, 73: 1-12.
- Piotroski, J. D., Wong, T. J., & Zhang, T. (2017). "Political bias in corporate news: the role of conglomeration reform in China," *The Journal of Law and Economics*, 60(1), 173–207.
- 邵軒磊 (Shao, H.) (2019). "機器學「習」：以文字探勘法探索習近平時期之大外宣戰略", 《中國大陸研究》 62 卷 4 期, pp.133-157.
- 邵軒磊 (Shao, H.) (2023). "政治人物重要講話之中文自然語言處理與意識形態體系模型", Available at SSRN: <https://ssrn.com/abstract=4324983>.
- Shirk, Susan L. (2018). "The Return to Personalistic Rule." *Journal of Democracy*, 29 (2): 22-36. Shirk, Susan. L. (2023). *Overreach: How China derailed its peaceful rise*. Oxford University Press.
- Sinkkonen, E. (2021). "Dynamic Dictators: Improving the Research Agenda on Autocratization and Authoritarian Resilience." *Democratization*, 28 (6): 1172-1190.
- The Economist (2016). "Chairman of everything," *The Economist*, 4 月 2nd, 2016.
- Watanabe, K. (2018). Newsmap: A semi-supervised approach to geographical news classification. *Digital Journalism*, 6(3), 294-309.

補足表 1. メディア別記事数(ダウンロード時点データ)

s	Media	N		Media	N
1	人民日报&人民网	6,764	31	学习时报	2
2	新华社&新华网	5,438	32	海南日报	2
3	求是	80	33	海外网	2
4	NA	76	34	贵州日报	2

5	央视网	16	35	中国军网	1
6	福建日报	12	36	中国网	1
7	中国新闻网	10	37	中央纪委国家监委网站	1
8	新京报	8	38	中青在线	1
9	解放军报	8	39	党建网	1
10	中国纪检监察报	7	40	内蒙古日报	1
11	浙江日报	6	41	北京晚报	1
12	中央纪委监察部网站	5	42	大众日报	1
13	甘肃日报	5	43	天津日报	1
14	中国共产党新闻网	4	44	学习小组	1
15	云南日报	4	45	宁夏日报	1
16	河北日报	4	46	山西日报	1
17	湖南日报	4	47	广西日报	1
18	解放日报	4	48	湖北日报	1
19	京华时报	3	49	监察部网站	1
20	南方日报	3	50	瞭望新闻周刊	1
21	新疆日报	3	51	经济日报	1
22	河南日报	3	52	郑州日报	1
23	陕西日报	3	53	青海日报	1
24	中国网络电视台	2			
25	光明日报	2			
26	北京日报	2			
27	北京青年报	2			
28	吉林日报	2			
29	四川日报	2			
30	国际在线	2			

出所：Xi-Database より筆者ら作成。

補足表 2. 分野別キーワード

分野	小分野	キーワード
A) 経済政策	a) 現状認識	新常态(新常态)、新发展阶段(新发展阶段)
	b) 構造改革	供給側改革(供給側改革/供給側结构性改革)、混合所有制(混合所有制)、決定的作用(决定性作用)、炭素中立(碳中和)、双循環(双循环)
	c) マクロ経済	マクロ調整(宏观调控)、減税(减税)、リスク解消(化解风险)、

		不動産(房地产)
	d) 政策課題	脱貧困(脱贫)、就業問題(就业问题)、収入格差(收入差距)、高齢化(老龄化)、城鎮化(城镇化)
	e) 競争政策	公平競争(公平竞争), 反独占(反垄断)、平台独占(平台垄断)
	f) 科学技術イノベーション政策	中国製造 2025(中国制造 2025)、断絶的技術(颠覆性技术)、軍民融合(军民融合)、大衆創業(大众创业)、戦略的新興産業(战略性新兴产业)、基礎研究(基础研究)
B) 内政	a) スローガン	中国夢(中国梦)、共同富裕(共同富裕)、偉大な復興(伟大复兴)、二つの百年(两个一百年/两个百年), 小康社会(小康社会)
	b) 反腐敗運動	反腐(反腐)、純潔(纯洁)
	c) 法治制度	法治(法治)、憲政(宪政)
	d) デジタル統治	デジタル化(数字化)、ネットワーク安全(网络安全)、データ安全(数据安全)、ネットワーク主権(网络主权)
	e) コロナ対策	動態清零(动态清零)
C) 外交	a) 認識	大有作為(大有作为)、韜光養晦(韬光养晦)、戦略機遇期(战略机遇期)
	b) 外交方針	新型国際関係(新型国际关系)、新型大国関係(新型大国关系)、周辺外交(周边外交)、核心的利益(核心利益)
	c) 独自構想	一帯一路(一带一路)、運命共同体(命运共同体)、AIIB(亚洲基础设施投资银行/亚投行)
	d) 国際機関	国際連合(联合国), 世界銀行(世界银行), WTO(世界貿易組織), IMF(国際貨物基金)
	e) 多国間枠組み	SCO(上海合作組織), BRICS(金砖国家)、G20(二十国集團), G7(七国集團)
	f) 経済連携協定	RCEP(区域全面經濟伙伴关系協定/RCEP), TPP(跨太平洋伙伴/跨太平洋战略經濟伙伴关系協定/TPP)
	g) 米中摩擦	一国主義(单边主义)、保護主義(保护主义)、いじめ主義(霸凌主义)
D) 安全保障	a) 概念	新型挙国体制(新型举国体制)、総体国家安全観(总体国家安全观)
	b) 領域	食糧安全(糧食安全)、エネルギー安全(能源安全)、ネットワーク安全(网络安全)、データ安全(数据安全)、経済安全(經濟安全)、国家安全(国家安全)

注 1: カッコ内が検索で用いた中国語である。例えば供給側改革の場合、複数の語彙で検索

している。これらキーワードの集計結果は伊藤(2023)を参照。

注2: このほかに特定の政策アジェンダを含意する地域名として香港、新疆、台湾を別途集計している。

注3: キーワードの選定に際しては、基礎資料として、①五か年計画(第13次、第14次)、②各年の政府工作報告、③毎年12月開催の中央経済工作会議、④習近平政権1-2期で2回開催された中央外事工作会議(2014, 2018)、⑤『求是』に掲載される重要政策キーワードに関する論文(習近平を著者とするものを含む)、⑥『習近平国政運営を語る(談国理政)』、⑦その他の資料・報道を参照し、政策キーワードの候補をリストアップした。

出所: 筆者ら作成。

補足表3. 各キーワードの登場頻度

	(A) 総記事数	(B) 通期比率	(C) 第一期比率	(D) 第二期比率	(E) 比率最大月
偉大な復興	1,929	19.8%	19.4%	20.3%	2017年10月
一帯一路	1,745	17.9%	12.8%	23.3%	2019年4月
中国夢	1,672	17.2%	20.7%	13.5%	2018年2月
運命共同体	1,562	16.1%	9.4%	23.1%	2018年7月
小康社会	1,310	13.5%	13.6%	13.4%	2012年11月
法治	1,214	12.5%	14.3%	10.6%	2015年2月
国際連合	1,186	12.2%	12.4%	12.0%	2015年9月
脱貧困	982	10.1%	5.9%	14.5%	2021年2月
二つの百年	874	9.0%	9.6%	8.3%	2016年12月
核心的利益	780	8.0%	8.3%	7.7%	2015年4月
G20	553	5.7%	7.1%	4.2%	2016年9月
国家安全	546	5.6%	5.6%	5.6%	2022年10月
反腐	450	4.6%	6.1%	3.1%	2015年1月
供給側改革	414	4.3%	3.7%	4.9%	2017年3月
BRICS	406	4.2%	5.4%	2.9%	2017年9月
新疆	404	4.2%	4.6%	3.6%	2017年2月
香港	400	4.1%	3.5%	4.8%	2022年7月
SCO	398	4.1%	4.8%	3.4%	2018年6月
保護主義	364	3.7%	2.9%	4.6%	2017年1月
城鎮化	342	3.5%	5.2%	1.7%	2013年12月
新型国際関係	331	3.4%	2.9%	4.0%	2018年6月
台湾	319	3.3%	3.3%	3.3%	2021年2月
糧食安全	289	3.0%	2.9%	3.0%	2013年12月

共同富裕	283	2.9%	1.9%	3.9%	2022年8月
AIIB	278	2.9%	5.1%	0.5%	2014年11月
新常态	253	2.6%	4.7%	0.4%	2015年3月
純潔	244	2.5%	2.8%	2.2%	2015年8月
一国主義	225	2.3%	0.2%	4.6%	2020年11月
WTO	211	2.2%	1.3%	3.1%	2018年11月
网络安全	186	1.9%	2.5%	1.3%	2014年5月
新发展阶段	182	1.9%	0.0%	3.8%	2021年2月
决定的作用	162	1.7%	1.9%	1.4%	2013年12月
新型大国関係	143	1.5%	2.8%	0.0%	2015年9月
マクロ調整	139	1.4%	1.7%	1.1%	2012年12月
軍民融合	139	1.4%	1.7%	1.1%	2018年3月
デジタル化	128	1.3%	0.4%	2.3%	2020年8月
エネルギー安全	124	1.3%	1.2%	1.3%	2022年6月
周辺外交	123	1.3%	2.0%	0.5%	2015年11月
双循環	114	1.2%	0.0%	2.4%	2020年11月
公平競争	113	1.2%	0.8%	1.5%	2020年12月
IMF	106	1.1%	1.4%	0.8%	2015年11月
基礎研究	106	1.1%	0.5%	1.7%	2021年5月
戦略機遇期	103	1.1%	1.4%	0.7%	2012年11月
戦略的新興産業	102	1.0%	1.1%	1.0%	2020年4月
炭素中立	96	1.0%	0.0%	2.0%	2021年7月
総体国家安全観	87	0.9%	0.7%	1.1%	2017年10月
大有作為	87	0.9%	1.1%	0.7%	2021年5月
RCEP	84	0.9%	0.5%	1.2%	2020年11月
不動産	80	0.8%	0.8%	0.8%	2012年12月
世界銀行	68	0.7%	0.8%	0.6%	2015年10月
減税	50	0.5%	0.2%	0.8%	2020年2月
高齢化	49	0.5%	0.5%	0.5%	2016年8月
大衆創業	46	0.5%	0.7%	0.3%	2018年1月
データ安全	44	0.5%	0.1%	0.8%	2020年11月
リスク解消	44	0.5%	0.2%	0.7%	2022年4月
経済安全	42	0.4%	0.3%	0.5%	2017年2月
収入格差	42	0.4%	0.4%	0.5%	2018年1月
新型拳国体制	36	0.4%	0.0%	0.7%	2021年5月
動態清零	31	0.3%	0.0%	0.7%	2022年7月

反独占	30	0.3%	0.1%	0.5%	2020年12月
断絶的技術	28	0.3%	0.2%	0.4%	2018年5月
就業問題	26	0.3%	0.2%	0.4%	2016年4月
G7	25	0.3%	0.4%	0.1%	2017年9月
中国製造 2025	21	0.2%	0.4%	0.0%	2015年10月
混合所有制	19	0.2%	0.3%	0.1%	2014年3月
TPP	18	0.2%	0.0%	0.4%	2021年11月
ネットワーク主権	18	0.2%	0.2%	0.1%	2017年12月
憲政	6	0.1%	0.0%	0.1%	2019年2月
プラットフォーム独占	5	0.1%	0.0%	0.1%	2021年10月
いじめ主義	2	0.0%	0.0%	0.0%	2019年12月
韜光養晦	2	0.0%	0.0%	0.0%	2013年3月

注：通期登場比率で順位付け。第一期は2012年11月から2017年10月、第二期は2017年11月から2022年10月としている。第19回中国共産党全国代表大会は2017年10月18-24日に開催されているが、ここでは2017年11月以降を第二期として集計している。

出所：Xi-Databaseより筆者ら作成。