

# 構造トピックモデルを用いた レコード・チャイナの分析

呉 東文

(ディップ株式会社商品開発本部データサイエンティスト)

2023年12月13日

\*本論考は計量政治学とデータを用いた分析で導き出された筆者の個人の意見であって、所属する組織の見解を示すものではありません。

## 要旨

本稿では、レコード・チャイナ(Record China)のテキストデータを使用して、運営者がどのような言説を広めようとしているのかを分析する。分析を通じて、レコード・チャイナというサイト名を持っているものの、日韓関係の負の側面に関連する記事が多いことが判明した。具体的には、慰安婦問題や旭日旗をめぐる韓国の日本に対する批判や、韓国側の行動に関連する記事が見られる。また、中国関連の報道に関しては、中国が国際協調で影響力を発揮していることや、中国の経済力と技術力の高さを強調する内容が目立つ。これはいずれも中国がアメリカと対等な関係にあるか、もしくはアメリカを越えていることを見せようとしている意図が背後にあると思われる。これはいずれも一橋大学の市原麻衣子教授の過去の別手法・別データを使った研究でも実証されたことであるため、信憑性は高いといえよう。ただし、ここで強調しておきたいのは、本記事は二年間かつ単独のメディアのみを使った分析であるため、エビデンスの強さに限界がある。例えば、分析で判明した傾向は果たしてレコード・チャイナにしか見られない傾向なのか、それとも他の日本語のメディアにも見られるような言説なのかがわからない。これは、今後データを追加することでよりはっきり識別できる。

## はじめに

本稿では、一橋大学の市原麻衣子教授から提供を受けたスクレイピングデータである 2021 年と 2022 年のレコード・チャイナ(Record China)のテキストデータを利用する。そして [Ichihara\(2020a\)](#)が指摘するようにレコード・チャイナのサイト運営に影響を与えていると思われる中国が、日本でどのような言説を広めようとしているのかを分析する。

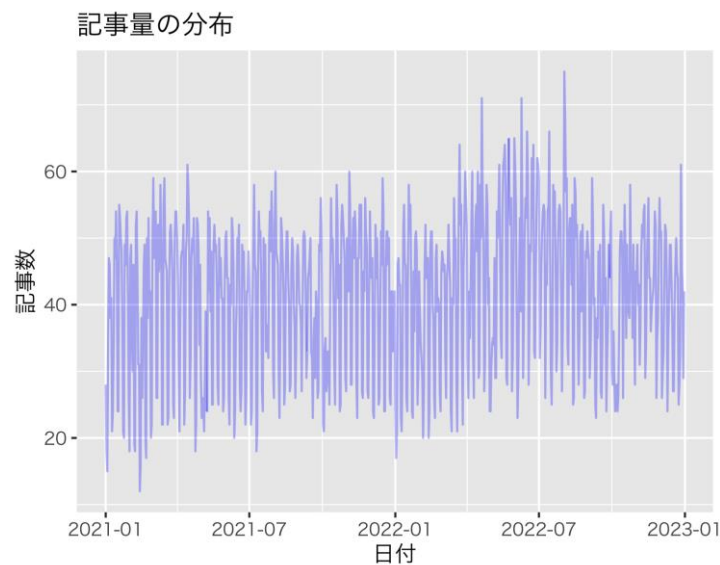
具体的な分析手法としては、政治学のテキスト・アズ・データ(text as data)の手法である構造トピックモデル(Structural Topic Model: STM)を採用する。STM を利用することで、配信元や配信時期などの変数がどのようにトピックの言及されやすさに影響するかを可視化できるため、サイト運営戦略のより緻密な分析が可能になる。

ただし、トピックモデルのような大規模ベイズモデルの結果を全て解釈しようとすると、本稿にまとめられない分量になるため、本稿ではレコード・チャイナが日韓関係の悪さと中国の経済、技術の優位性を強調する側面に焦点を当てて分析を進めたい。

## データ説明

市原麻衣子教授から、スクレイピングで取得した約 30,000 件のデータを受領した。これは 2021~2022 年にレコード・チャイナから出版された全記事である。記事量の推移を確認すると、少なくとも 2021 年と 2022 年の 2 年間では毎日安定して 20 件から 60 件ほどの記事を掲載していることが読み取れ、レコード・チャイナは定常的に運用されていることが確認できる。

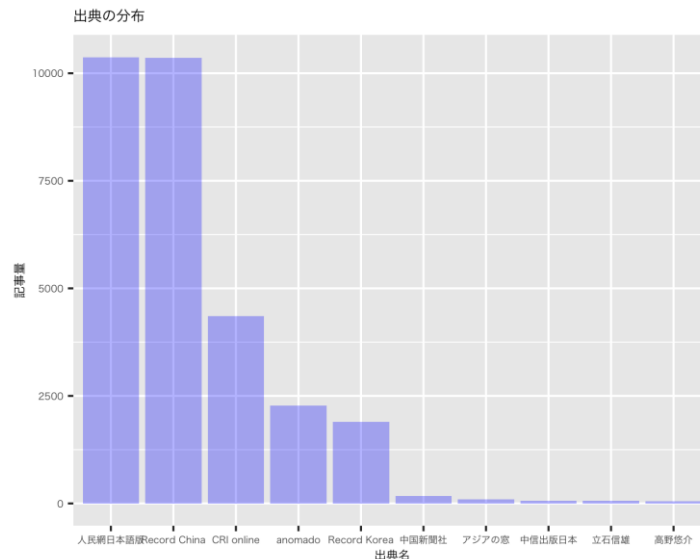
図1 レコード・チャイナ記事量の分布



出典: 著者作成

レコード・チャイナの記事は、別の媒体から出版された記事を転載しているものが多い。そこで次に、レコード・チャイナのサイト上で、記事のタイトルの下に記載された出典の分布を確認する。ただ、個人が執筆した記事などもあり、全てを可視化するのは困難であるため、記事量が最も多い 10 出典に絞って可視化する。

図 2 出典の分布



出典: 著者作成

図からもわかるように、スクレイピングで取得した約 30,000 件の記事データの出典の大半をレコード・チャイナ自身の記事と人民網日本語版から転載された記事が占めているため、本稿では出典がレコード・チャイナか人民網日本語版の記事に絞って分析する。

## 分析手法

### LDA の説明

前述のように、本記事では STM を利用して分析を行うが、そもそもトピックモデルとは何かを説明する。

トピックモデルは、基本的には全てデイビット・ブレイ(David Blei)らが 2003 年の論文(Blei et al. 2003)で提案した潜在的ディリクレ配分法(Latent dirichlet allocation: LDA)というテキストデータ分析の手法から派生したものであり、根底にある概念は概ね同じである。そのため、本記事はまず LDA の基本的な考えを説明する。

LDA では、文章は複数のトピックの割合で表現でき、かつトピックは異なる単語の出現確率を持つと仮定する。そこで、モデルは自動で各文章のトピックの割合と各トピックの単語の出現確率を推定・学習する。たとえば、とある政治学の論文は 40%の割合で方法論の話をしているが、残りの

60%の割合で歴史と背景を説明すると、方法論のトピックが 40%で歴史のトピックが 60%という形で表現される。トピックでいうと、方法論トピックは「有意性」、「識別戦略」、「事前分布」などの単語が比較的高い出現確率を持つのにに対し、歴史のトピックは具体的な事件やアクターの名前が高い出現確率を持つと予想される。

文章とトピックの次に、文章中の単語モデルの生成について LDA はどんな仮定を置いているかを確認する。まず、文章のトピック割合から一つのトピックを抽出する。次に、その抽出されたトピックの単語の出現確率に従って単語を抽出するのである。

## STM の説明

以上の内容を踏まえて、次は本稿で利用する STM の説明に入る。

STM は、上述した LDA の基本構造の他に、普及共変量(prevalence covariate)と内容共変量(content covariate)という二つの共変量を持っている。普及共変量は、あるトピックの語られやすさに影響する変数のことである。例えば、農業に関心のある議員は農業トピックを他の議員より多く語っている事実を表すことができる。内容共変量というのは、内容のそもそもの語り方の違いを可視化するための変数のことである。わかりやすい事例でいうと、同じウクライナ・ロシア紛争のトピックでも、ウクライナメディアとロシアメディアは異なる内容を強調するであろう。しかし、内容共変量を指定しない STM や伝統的な LDA の場合、「ウクライナ版のウクライナ・ロシア紛争トピック」と「ロシア版のウクライナ・ロシア紛争トピック」が生じてしまう。内容共変量を STM に組み込むことで、モデルが「ウクライナ版のウクライナ・ロシア紛争トピック」と「ロシア版のウクライナ・ロシア紛争トピック」を一つの「ウクライナ・ロシア紛争トピック」に統合して、その中のウクライナとロシアの語り方の違いを可視化することができる。

## データ説明

ここではデータを STM に入れて分析を行う。前述したように、本記事の分析では、市原麻衣子教授から提供を受けたデータの中で、レコード・チャイナの記者が執筆したものと、人民網日本語版の記事を転載しているものを対象とする。日本語の処理に関しては、ひらがな、カタカナ、漢字、ローマ字以外のものは全てスペースに置き換える。次に、RMeCab 経由で mecab を利用して名詞だけを残し、R のテキストデータ処理パッケージ `quanteda` でトークン化や文書行列の作成を行い、最後に「の」とその他の出現頻度が特に高い単語と特に低い単語を除去して、STM のデータとして投入する。

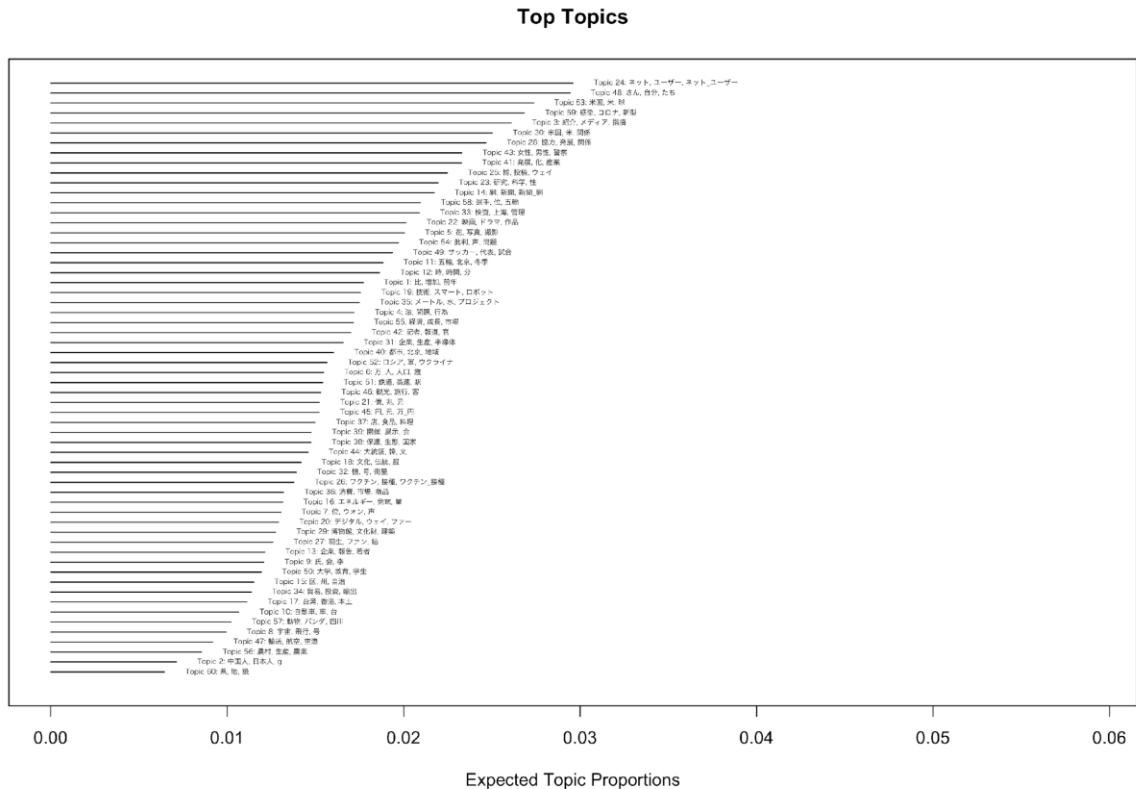
STM を利用する際には、トピック数を指定する必要があるため、本分析では 60 に指定する。また、上述した普及共変量に関しては出典名(レコード・チャイナか人民網日本語版)と月を設定した。内容共変量の指定はない。他のパラメータ等は全て STM のパッケージのデフォルト設定に準拠する。

## 全体像

まずは、推定された結果の全体像から確認する。図 3 は 60 のトピックのそれぞれの代表的な単

語と、それぞれが全体のデータセットの中で占めている割合を示している。

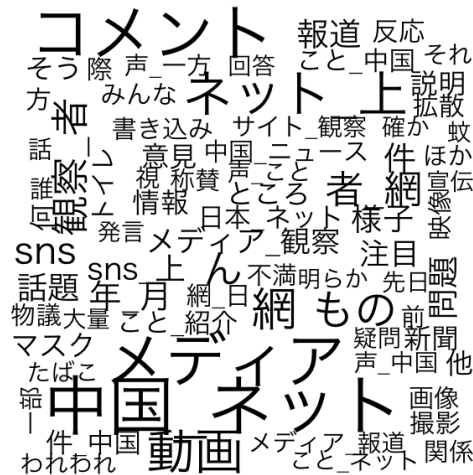
図3 STMのトピック分布



出典: 著者作成

この図よりまず指摘できることは、市原麻衣子教授から提供を受けたレコード・チャイナのデータの中で割合が一番高いトピックはトピック 24 で、内容としてはインターネット上の声に関連するものである。下記の図 4 は、トピック 24 の詳細を確認するために、トピック 24 の代表的な単語をワードクラウド形式で表したものである。

図4 トピック24(ネット上の声トピック)のワードクラウド



出典: 著者作成

このワードクラウド内の文字が大きければ大きいほど、このトピックにおける代表的な単語であることを示す。図4より、確かに「ネット\_ユーザー」や「ネット」「声」などの単語が確認できる。レコード・チャイナがインターネット上の声を強調するのは、運営者が閲覧者の親近感を高め、言説を広めやすくする狙いがあると推測される。

このモデルでは60トピックを推定しており、これらの中には米中関係や台湾関連のトピックなど、多種多様な内容が含まれている。なお、全てのトピックについて深掘りするのには紙幅の関係でできないため、「はじめに」で述べたように、読者が特に関心を持ちそうな日韓関係関連と中国関連のトピックについて説明する。

### 日韓関係関連の言説

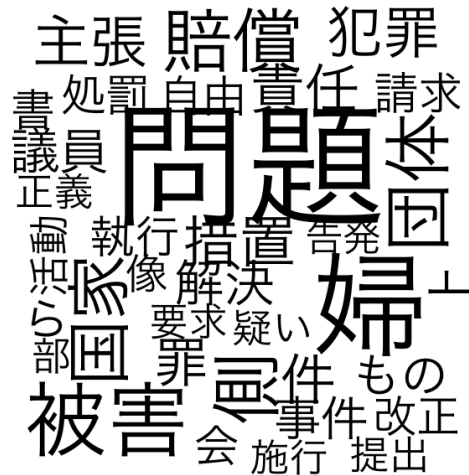
まず、興味深いことに、レコード・チャイナはそのサイト名にもかかわらず、慰安婦問題と旭日旗に関連する記事を掲載している。さらに図3からも読み取れるように、慰安婦問題に関連するトピック4も旭日旗に関連するトピック54もデータセットの中で割合が少ない訳でもない。したがって、レコード・チャイナの運営者は、意図的に慰安婦問題や旭日旗関連問題に代表されるような、日韓関係問題を強調しようとしているという仮説が立てられる。

ここで、慰安婦問題関連と旭日旗関連の内容はそれぞれどのようなものかを詳しく確認する。

### 慰安婦問題

まず、内容を視覚的に把握するため、慰安婦問題に関連するトピック4のワードクラウドを作成した。

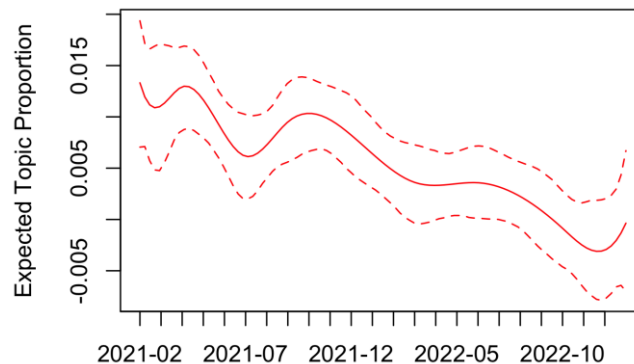
図5 トピック4(慰安婦問題トピック)のワードクラウド



出典:著者作成

普及共変量を確認すると、人民網日本語版の係数が-0.024061かつ0.1%有意になっているため、人民網日本語版よりレコード・チャイナが取り上げる傾向が高いトピックだと判断できる。月も普及共変量としてモデルに入れているため、毎月のトピック4の取り上げられやすさは下記のように可視化できる。

図6 トピック4(慰安婦問題トピック)の取り上げられやすさの推移



出典:著者作成

図6からは、トピック4の取り上げられやすさ、すなわち慰安婦問題が取り上げられる傾向は下降傾向にあることが読み取れる。

次に、実際のレコード・チャイナの[記事](#)の一部を確認する。

2021年6月15日、韓国・中央日報によると、ソウル中央地裁が日本政府に元慰安婦らへの賠償を命じた判決をめぐり、地裁が日本政府に対し韓国内に保有する資産の目録を開示するよう命じた。記事によると、元慰安婦12人が起こした今年1月の損害賠償請求訴訟で、キム・ジョンゴン裁判長は1人当たり1億ウォン(約980万円)支払うよう日本政府に命じた。国家には他国の裁判権が及ばないとする国際法上の「主権免除」を主張する日本政府は控訴せず、判決は確定。慰安婦らは強制執行のため同地裁に「財産明示申請」を行っていた。(Record China 2021a)

また、レコード・チャイナには次のような記事も掲載されている。

2021年8月27日、韓国・デジタルタイムズによると、元慰安婦関連団体に対する「事実摘示型の名誉毀損(きそん)」を厳しく処罰する内容が盛り込まれ物議を醸していた、いわゆる「尹美香(ユン・ミヒャン)法」が撤回された。元慰安婦支援団体「日本軍性奴隷制問題解決のための正義記憶連帯(正義連)」の前理事長である尹美香(ユン・ミヒャン)議員(無所属)は13日、与党「共に民主党」の印在謹(イン・ジェグン)議員らと「日帝下の日本軍慰安婦被害者に対する保護・支援および記念事業などに関する法律の一部改正法律案」を共同発議した。(Record China 2021b)

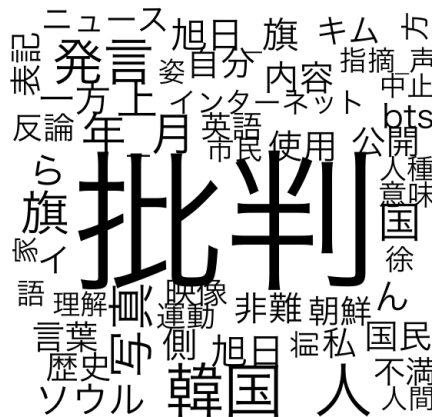
このように、元慰安婦の話や元慰安婦保護と支援関連の内容も含まれている。

## 旭日旗

旭日旗に関連するトピック 54 についても、まず内容を視覚的に把握するためのワードクラウドを作成した。



図7 トピック 54(旭日旗トピック)のワードクラウド

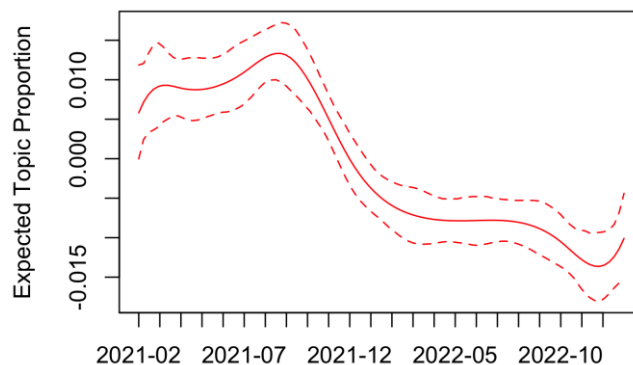


出典: 著者作成

普及共変量を確認すると、人民網日本語版の係数が-0.033791 かつ 0.1%有意になっている。つまり、慰安婦トピックと同様に、旭日旗トピックは人民網日本語版よりレコード・チャイナが取り上げがちなトピックだといえる。サイト名からしていかにも中国専門メディアのようなレコード・チャイナが、日韓関係の悪さを強調する内容を人民網日本語版よりも多く取り上げている事実は興味深い。もっとも、何らかの理由で日韓関係の悪さを強調する人民網日本語版の記事の転載を控えている可能性があるため、転載されていない人民網日本語版の記事を含めても同じ傾向が出るかに関しては追加分析に値する。

次に、トピック 54 の毎月の取り上げられやすさを可視化する。

図8 トピック 54(旭日旗トピック)の取り上げられやすさの推移



出典: 著者作成

慰安婦トピックと同様、全体的な取り上げやすさは下降傾向にあるといえよう。

次に、旭日旗トピックの実際の記事の内容の一部を確認するとこのように、大学教授の発言からの引用ではあるものの、「戦犯旗」という強い表現を使っている。

2021年10月15日、英国の世界的ロックバンド、レッド・ツェッペリンの公式ウェブサイトに「旭日旗」が使用されていることに対し、韓国の広報活動をする徐敬徳(ソ・ギョンドク)誠信女子大学教授が抗議のメールを送ったことを SNS で明らかにした。徐教授によると、レッド・ツェッペリンの公式ウェブサイトでは現在、1971年の日本初公演から50周年を迎えたことを記念するTシャツが販売されているが、ウェブサイトの背景とTシャツのデザインに旭日旗が使用されている。これに対し徐教授は「旭日旗は太平洋戦争時に日本の軍隊が使用した軍旗で、『戦犯旗』にあたる」「戦犯旗をウェブサイトやTシャツデザインに使用することは多くのアジアのファンを傷つける行為」などと説明し、「ただちに修正し、世界的ポップスターとして世界のファンに向けて良い前例を作してほしい」と求める内容の抗議メールを送ったという。(Record China 2021c)

また、次の記事でもやはり「戦犯旗」として旭日旗を表現している。

2021年4月29日、韓国・スポーツウエイによると、韓国の広報活動をする誠信女子大学の徐敬徳(ソ・ギョンドク)教授がカナダの歌手ジャスティン・ビーバーの「旭日旗衣装」を批判した。記事によると、ジャスティンは今月9日に放送された「ミュージックステーション(Mステ)」に出演した際、日の出のモチーフがデザインされた米ファッションブランド「ERL」のダウンジャケットを着用していたが、これが韓国で「旭日旗を連想させる」として物議を醸していた。これを受け徐教授は29日、自身のSNSで「ジャスティンのSNSアカウントと所属事務所に『旭日旗はハーケンクロイツと同じ意味を持つ戦犯旗だ。西洋では知らない人が多い』と説明する内容の抗議メールを送った」と明らかにした。また「今回のことをきっかけに旭日旗の歴史的意味を学び、二度と繰り返さないでほしい」とし、「アジアのファンに心から謝罪せよ」と求めたという。(Record China 2021d)

市原麻衣子は、中国が日本において韓国に対するネガティブな言説を広めようとしていることを指摘している(Ichihara 2020b)。異なる方法論とデータを利用する本稿でも同様の現象を発見できたことは、かなり信憑性の高い仮説であることを示している。

### 中国関連の言説

手元には2021年から2022年までのデータしかないため、時期の問題の可能性もあるものの、中国関連の報道は、日韓関係関連の報道とは対照的である。中国を取り上げたトピックは、よりソフトな観光や各国の協調を促進する自国の役割や、経済及び技術の発展を取り上げることが多い。そのため、日中関係の悪さや課題などはほぼ取り上げられていない。



国際連合安全保障理事会の今月の輪番議長国を務める中国の提案より、安保理は 7 日にオンライン方式で「国際平和と安全の維持：多国間主義と国連中心の国際体制を守る」をテーマにしたハイレベル会合を開催した。同会合では中国の王毅国務委員兼外交部長（外相）が議長を務めた。新華社が伝えた。王部長は、「昨年 9 月、習近平国家主席が国連創設 75 周年を記念した一連のハイレベル会合に出席し、多国間主義のについて踏み込んだ解釈を行い、多国間主義の道を堅持すべきだと明確に指摘した」と指摘。また、「国連は多国間主義の旗印だ。多国間主義を堅持するには、『国連憲章』の趣旨と原則を礎石とした国際関係の基本ルールを維持し、国連の権威と地位を維持し、国際問題における国連の中核的な役割を維持することが必要だ」と述べた。(Record China 2021e)

次の記事にも同じような言説が見られ、特にデカップリングの逆流の阻止や多国間主義の擁護者、世界の発展への貢献者などの用語も使っている。したがって、アメリカに代わって国際強調を維持する大国になりたい、もしくはすでになっていることを表そうとしていると解釈できる。

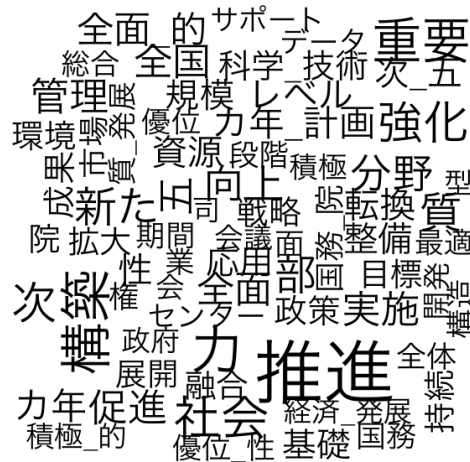
中国の王毅(ワン・イー)国務委員兼外交部長(外相)は 21 日、ドイツのマース外相とテレビ会談を行った。王部長は、「昨年以來、中独関係は全体的に安定を保っている。包括的戦略パートナーとして、中独は常に相互尊重という重要な原則と貴重なノウハウをしっかりと押さえ、対話と協力という主旋律を堅持し、協力・互惠・ウィンウィンという積極的なメッセージを対外発信している。双方は共にいわゆる『デカップリング』の逆流を阻止し、多国間主義の擁護者、世界の発展への貢献者であるべきだ」とした。マース外相は、「ドイツは対中関係を非常に重視しており、各分野で対話を増進し、実務協力を強化し、グローバルな問題や試練に共同で対処することを望んでいる」とした。(Record China 2021f)

### 経済と技術の発展

経済発展を強調するトピックはトピック 41(中国のイノベーション政策トピック)、トピック 23(中国の医療技術トピック)、トピック 19(中国のイノベーショントピック)、トピック 55(中国とアジアの経済情勢トピック)など複数あるため、紙幅の制約で他のトピックのように緻密に分析結果を説明することができない。しかし、まず強調しておきたいのは、トピック 41、トピック 23、トピック 19、トピック 55 はいずれも人民網日本語版の係数が正で、人民網日本語版はレコード・チャイナより中国の経済と技術の発展を取り上げやすいことである。ここからもサイト運営者は何かしらの基準と目的でレコード・チャイナと人民網日本語版の使い分けをしていることがわかる。

まずはトピック 41 から確認する。

図 11 トピック 41(中国のイノベーション政策トピック)のワードクラウド



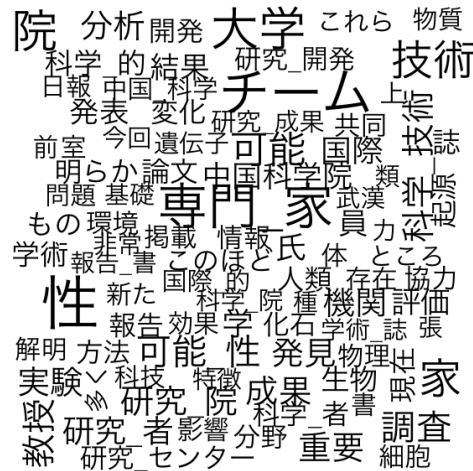
出典: 著者作成

実際の内容を合わせて見ると、これは中国政府のイノベーション関連政策の方向性などの[記事](#)がまとめられたトピックだとわかる。

中国共産党中央委員会と国務院の通達した「知的財産権強国建設綱要(2021-2035年)」が 22 日に発表された。これにより、知的財産権強国の建設加速への方針が整えられた。新華社が伝えた。綱要は「中国の特色ある、世界水準の知的財産権強国の建設は、国家のコア・コンピタンスを高め、高水準の対外開放を拡大し、より質が高く、より効率的で、より公正で、より持続可能で、より安全な発展を実現し、素晴らしい生活への国民の高まるニーズを満たす上で、重要な意義を持つ」と指摘。「2025 年までに、知的財産権強国の建設において著しい成果を収め、知的財産権の保護をより厳格にし、社会的満足度を高い水準にして維持し、知的財産権の市場価値を一層明らかにし、ブランド競争力を大幅に高める。2035 年までに、知的財産権の総合競争力において世界トップレベルになり、中国の特色ある、世界水準の知的財産権強国を基本的に完成する」とした。(Record China 2021g)

次に、トピック 23 を確認すると、このトピックは遺伝子等のキーワードが入っていることから、医療系の技術に関連するトピックであることがわかる。

図 12 トピック 23(中国の医療技術トピック)のワードクラウド



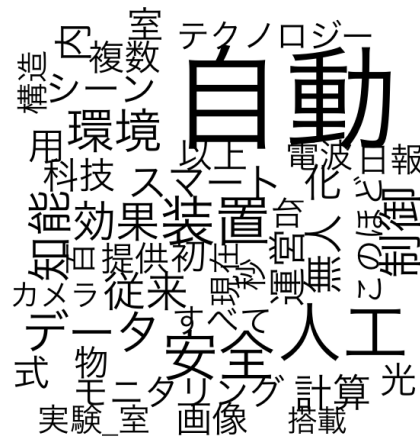
出典:著者作成

実際の記事を合わせて見ると、本トピックは中国政府のイノベーション関連政策の方向性などの記事がまとめられたトピックだとわかる。筆者は医学の専門家ではないので詳しい評価は避けるが、実トピック 23 に関連する記事を確認すると、高度な技術の紹介が印象的である。

武漢大学が 10 日に明らかにしたところによると、同大の中南病院の研究者は、あるリソソーム関連タンパク質(TMBIM1)が脂肪の生成に対して鍵となるマイナス調整の役割を持つことを発見した。これは肥満関連の代謝の乱れを治療する潜在的な分子標的だ。同研究は 8 日、代謝分野のトップレベルの国際学術誌「Cell Metabolism」にオンライン掲載された。科技日報が伝えた。同研究は生物情報学的手段によりマウスとヒトの前駆脂肪細胞分化のトランスクリプトームデータを分析し、TMBIM1 の脂肪細胞分化における発現の顕著な変化を発見した。研究は、TMBIM1 が体外前駆脂肪細胞の分化モデルの中で、脂肪細胞の分化を抑制する機能を発揮することを証明した。(Record China 2021h)

次は中国の人工知能などのイノベーションを取り上げるトピック 19 を紹介する。

図 13 トピック 19(中国のイノベーショントピック)のワードクラウド



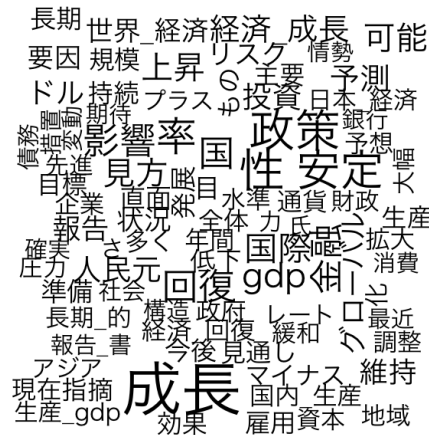
出典: 著者作成

[以下の記事内容](#)からわかるように、このトピックは主に中国が有する先進的技術をアピールすることが目的だと思われる。

ジージーという音と共に、「虹拓」や「中国光谷」などの漢字と図形が空中に浮かび上がり、肉眼で確認でき、手で触れられる。これは 19 日、武漢東湖ハイテク開発区の虹拓超高速レーザー共同実験室で展開された SF のようなシーンだ。同実験室の曹祥東室長によると、チームが新たに研究・製造した高エネルギー・高ピーク出力のフェムト秒レーザー装置のピーク出力密度は 100 万 MW 級に上り、「最も速い光」を束にして集め、空気を明るくすることができる。科技日報が伝えた。(Record China 2021i)

最後に、中国とアジアの経済情勢の関連するトピックを確認する。

図 14 トピック 55(中国とアジアの経済情勢トピック)のワードクラウド



出典:著者作成

トピック 55 に関連する記事の内容を確認すると、アメリカとヨーロッパの景気の悪化とアジア地域の経済的な影響力を強調する内容もあれば数字を紹介する内容もある。

経済協力開発機構(OECD)は22日に発表した経済見通しの中で、グローバル経済成長率について、2022年は3.1%になり、23年は2.2%に鈍化し、24年は2.7%になるとの予測を打ち出した。新華社が伝えた。同見通しによれば、ウクライナ危機により発生したエネルギー問題がインフレを刺激し続け、世界的なリスクを増大させていることから、今後1年間はグローバル経済の成長がさらに鈍化することが予測されるという。同見通しは、「アジアの主要な新興市場国・地域は23年に世界の国内総生産(GDP)の成長の4分の3近くを占めるようになる。米国と欧州の経済成長率は急激に鈍化しつつある。米国経済の今年の成長率は1.8%、来年は0.5%に低下する。ユーロ圏経済の今年の成長率は3.3%、来年は0.5%に低下し、英国経済の今年の成長率は4.4%、来年はマイナス成長に転じ、マイナス0.4%となる見込みだ」と予測している。(Record China 2022)

また、以下のようなトピック 55 に関連する記事も掲載されている。

国際通貨基金(IMF)は6日、最新の「世界経済見通し(WEO)」を発表した。新華社が伝えた。2021年の中国経済は8.4%成長すると予測し、1月の予測値よりも0.3ポイント上方修正した。(Record China 2021j)

前述した市原麻衣子の研究においても、中国がメディアを利用して中国の良さを広めようとしている現象が指摘されている(Ichihara 2020)。したがって、韓国関連だけでなく中国関連で



も、本稿の分析結果と市原麻衣子の研究結果が類似している。中国が日本で日韓関係の負の側面や中国の良さに関連する言説を広めようとしているのは、特定の期間とメディアを越えた一般的な現象である可能性が高いと判断できる。

## 結論

本記事では、STM を使用してレコード・チャイナの記事を分析することで、以下の二つの現象を確認することができた。第一に、レコード・チャイナは一見、中国を中心とする情報を紹介するサイトのように思われる一方で、実際は韓国関連、特に日韓関係の負の側面を取り上げている。二つ目の現象として結論で繰り返し強調したいのは、レコード・チャイナ、特に出典が人民網日本語版である記事は中国の良さを強調する一方、日中関係の問題点を日韓関係の問題点のように紹介していない。これはサイト運営者の何らかの戦略を反映している可能性が高い。

そして本稿の分析の限界もここで強調したい。まず、市原麻衣子教授から得たデータは 2021 年と 2022 年の二年間のみのものであるため、STM による分析で明らかになったことは本当にレコード・チャイナの傾向なのか、それともこの二年間の偶然なのかを判断することができない。この問題を解決するために、まずデータの期間を追加することで、過去日中関係が極めて悪い時期にレコード・チャイナは日中関係についてどのように報道したのかを分析することができる。しかしながら、[Ichihara\(2020a\)](#)による 2019 年のデータを使用した分析でも同様の結果が検証されたため、ある程度の妥当性があるといえよう。他にも、日本の大手全国紙など他の情報を追加することで、今後はより緻密な分析が期待される。

参考文献

- Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003). “Latent dirichlet allocation.” *Journal of Machine Learning Research*, 3, 993-1022.
- Ichihara, Maiko (2020a). “Is Japan Immune from China’s Media Influence Operations?” *The Diplomat*. (<https://thediplomat.com/2020/12/is-japan-immune-from-chinas-media-influence-operations/> 2023年11月24日最終閲覧)
- (2020b). “Country Case 1: Japan Influence Activities of Domestic Actors on the Internet: Disinformation and Information Manipulation in Japan.” *[Special Report] Social Media, Disinformation, and Democracy in Asia: Country Cases*. Asia Democracy Research Network, 1-29.
- Record China (2021a). 「韓国裁判所、対日本政府訴訟でまた『正反対の決定』=韓国ネット『コロコロ変わりすぎ』『日本が従うと思う?』」(<https://www.recordchina.co.jp/b878032-s25-c100-d0194.html> 2023年11月9日最終閲覧)
- (2021b). 「批判続出の”元慰安婦保護法”が撤回、発議した韓国議員は『残念だ』と落胆=韓国ネット『面の皮が厚い』」(<https://www.recordchina.co.jp/b878032-s25-c100-d0194.html> 2023年11月9日最終閲覧)
- (2021c). 「ジャスティン・ビーバー、エド・シーランに続き…世界的ロックバンドの『旭日旗』使用に韓国が憤慨」(<https://www.recordchina.co.jp/b883667-s25-c30-d0191.html> 2023年11月9日最終閲覧)
- (2021d). 「BTS 事務所に合流したジャスティン・ビーバーが“旭日旗”衣装!?韓国で批判の嵐、謝罪求めるメールも」(<https://www.recordchina.co.jp/b875719-s25-c30-d0191.html> 2023年11月9日最終閲覧)
- (2021e). 「国連安保理ハイレベル会合開催、王毅外交部長が議長務める」(<https://www.recordchina.co.jp/b876096-s6-c100-d0165.html> 2023年11月9日最終閲覧)
- (2021f). 「中独外相がテレビ会談、「デカップリング」の逆流阻止に尽力—中国メディア」(<https://www.recordchina.co.jp/b875342-s6-c100-d0189.html> 2023年11月9日最終閲覧)
- (2021g). 「知財権強国建設綱要『中国の特色ある、世界水準の知財権強国を2035年までにほぼ完成』—中国」(<https://www.recordchina.co.jp/b882759-s6-c100-d0189.html> 2023年11月9日最終閲覧)
- (2021h). 「中国の科学者、脂肪生成を抑える新メカニズムを発見—中国メディア」(<https://www.recordchina.co.jp/b877896-s6-c30-d0189.html> 2023年11月9日最終閲覧)
- (2021i). 「空気を明るくすることで3D画像を描くフェムト秒レーザー—中国」(<https://www.recordchina.co.jp/b898061-s6-c20-d0189.html> 2023年

11月9日最終閲覧)

—(2022).「OECD 予測「2023 年のグローバル経済成長率は 2.2%」—中国メディア」  
(<https://www.recordchina.co.jp/b904996-s6-c20-d0189.html> 2023 年  
11月9日最終閲覧)

—(2021j).「IMF、今年の中国経済は 8.4% 成長と予測—中国メディア」  
(<https://www.recordchina.co.jp/b874585-s6-c20-d0189.html> 2023 年  
11月9日最終閲覧)

Roberts, Margaret E., Brandon M. Stewart, and Edoardo M. Airolidi. “A model of text for experimentation in the social sciences.” *Journal of the American Statistical Association* 111.515 (2016): 988-1003.

#### 呉東文プロフィール

ディップ株式会社商品開発本部データサイエンティスト。国際・行政修士(専門職)。一橋大学国際・公共政策教育部国際・公共政策専攻(国際・行政コース)専門職学位課程修了後、東京で民間のデータサイエンティストとして就職し、現在に至る。専門はテキストアズデータ、因果推論、大規模ベイズ機械学習をはじめとする計量政治学と計量経済学。ブログは<https://qiita.com/Gotoubun taiwan>。