

基于深度学习技术的中国能源企业出海 英文新闻报道主题建模分析

张政汉, 郭佳然, 张 佳, 包宇璇

中国矿业大学(北京)文法学院, 北京

收稿日期: 2025年9月10日; 录用日期: 2025年9月30日; 发布日期: 2025年10月13日

摘 要

在全球化和能源转型背景下, 中国能源企业加速海外布局, 其出海成为国际舆论关注焦点。本研究采用基于深度学习的BERTopic主题建模方法, 结合框架理论对185篇国际英文媒体新闻报道进行分析, 系统揭示了中国能源企业出海报道的主题分布、地域特征与企业关注态势。研究发现, 国际媒体报道涵盖电动车竞争、油气合作、ESG治理等15个主题, 呈现传统与新兴议题并存格局; 地理分布上, 欧美舆论聚焦技术竞争与政策博弈, “一带一路”区域则侧重基建合作; 企业层面, 国际舆论焦点已从传统的油气企业显著转向新能源企业, 反映了中国能源企业在全局角色认知上的转变。本研究为中国能源企业应对国际舆论、优化传播策略提供了实证依据。

关键词

中国能源企业出海, 英文新闻报道, 主题建模, 深度学习, 框架理论

Topic Modeling Analysis of English News Reports on Chinese Energy Companies Expanding Overseas Based on Deep Learning Technology

Zhenghan Zhang, Jiaran Guo, Jia Zhang, Yuxuan Bao

School of Law and Humanities, China University of Mining and Technology (Beijing), Beijing

Received: September 10, 2025; accepted: September 30, 2025; published: October 13, 2025

文章引用: 张政汉, 郭佳然, 张佳, 包宇璇. 基于深度学习技术的中国能源企业出海英文新闻报道主题建模分析[J]. 统计学与应用, 2025, 14(10): 85-95. DOI: 10.12677/sa.2025.1410286

Abstract

Amid globalization and energy transition, Chinese energy companies are accelerating their overseas expansion, drawing significant attention from international discourse. This study utilizes BERTopic—a deep learning-based topic modeling method—integrated with framing theory to analyze 185 English-language news reports from international media on the global expansion of Chinese energy firms. It systematically uncovers the topic distribution, geographical patterns, and corporate focus within the coverage. Results identify 15 major topics, including electric vehicle competition, oil and gas cooperation, and ESG governance, reflecting a dual structure of both traditional and emerging issues. Geographically, coverage in Europe and North America emphasizes technology competition and policy challenges, whereas reports related to Belt and Road regions highlight infrastructure collaboration. At the corporate level, media attention has shifted markedly from traditional oil and gas enterprises to new energy players, signaling an evolution in the global perception of China's role—from resource acquirer to technology and market influencer. These findings provide an empirical basis for Chinese energy companies to shape more adaptive and effective international communication strategies.

Keywords

Chinese Energy Companies Going Global, English News Report, Topic Modeling, Deep Learning, Framing Theory

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在全球能源格局加速转型与中国“走出去”战略持续深化的双重背景下，中国能源企业的海外拓展不仅成为全球经济与能源治理的重要组成部分，也日益成为国际舆论关注的众多焦点议题之一。而各国英文媒体凭借其全球传播力与话语影响力，通过其新闻报道中蕴含的特定叙事框架建构对中国能源企业的认知与态度，深刻影响和塑造着中国企业海外经营的环境与形象。系统性研究英文新闻报道如何框架化中国能源企业出海实践，不仅具有学术层面的理论价值，更对察觉海外舆论与投资风险、优化企业国际传播策略具有重要的现实意义。

在此背景下，本研究运用先进的 BERTopic 主题建模文本语料挖掘技术，以框架理论(Framing Theory)为核心框架，致力于突破以往对中国能源企业出海偏重经济与国际关系的普遍论述。研究通过对 2002~2025 年间 185 篇主流英文媒体报道的主题建模与框架分析，系统考察英文媒体对中国能源企业出海的话语建构机制，旨在回答以下核心问题：(1) 英文媒体对中国能源企业出海的新闻报道有哪些核心主题？(2) 其报道呈现了什么样的地理与企业分布特征？(3) 这些媒体又采用了何种叙事框架，且这些框架如何潜移默化地塑造了中国能源企业的国际形象？

2. 文献综述

现有研究多聚焦于经济学与国际关系视角，探讨中国能源海外投资的驱动因素、市场风险与地缘政治影响[1]，或研究企业能源使用与出口行为的关联[2]。这些研究中虽也有关注意到欧美媒体的报道倾向[3]，

但普遍存在两点局限：其一，多停留于宏观判断或小样本案例，缺乏对海量媒体报道进行系统性量化分析与结构化框架识别，难以揭示舆论的整体构成与内在规律；其二，理论对话不足，未能深入运用传播学理论如框架理论来解构媒体文本背后的认知建构机制，致使对“媒体如何塑造中国能源形象”这一核心问题的解释受限。为弥补上述研究缺口，本研究引入框架理论(Framing Theory)作为核心理论框架，并基于深度学习技术的 BERTopic 主题建模方法，对主流英文媒体报道进行量化与质性分析，本研究将为理解中国能源企业当下所面临的国际舆论环境提供数据基础与理论支撑融合的新路径。

3. 理论框架与技术基础

3.1. 框架理论

框架理论(Framing Theory)源于社会学与传播学领域，由 Goffman (1974) [4]首次系统提出，用于解释个体如何通过特定的认知结构(即“框架”)来理解和解释社会现实。它指出，媒体报道并非客观中立的传递信息，而是通过特定的叙事策略和语言选择，塑造受众对事件的感知与态度。在新闻传播研究中，框架理论被广泛用于分析媒体如何通过选择、强调和排除某些信息，来建构公众对某一事件或议题的特定理解[5]。这种“框架”本质上是一种对公众认知的引导，影响人们如何看待中国能源企业出海的行为——是将其视为合作机遇还是竞争威胁。在国际舆论中，媒体常通过定义问题、归因因果、道德评价和提出解决方案等方式构建报道框架。本研究借助该理论，深入分析国际英文新闻报道如何通过特定叙事框架影响中国能源企业的国际形象与舆论环境。

3.2. BERTopic 主题建模

本研究主要采用的 BERTopic 是基于预训练语言模型(BERT)的主题建模方法[6]，相较于传统 LDA (潜在狄利克雷分配)模型，其核心优势在于能够捕捉文本的深层语义关联，而非仅依赖词频统计，该方法已在新闻文本企业趋势提取中得到有效应用[7]。其技术流程主要包括四个步骤：第一步是文本向量化，通过 BERT 预训练模型将每篇新闻报道转化为高维向量，向量的每个维度对应文本的特定语义特征，确保语义相似的文本在向量空间中距离更近；第二步是降维处理，采用 UMAP (Uniform Manifold Approximation and Projection)技术将高维向量降至低维空间，在保留核心语义信息的同时，减少后续聚类的计算复杂度；第三步是聚类分析，使用 HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise)算法对低维向量进行聚类，该算法无需预先设定聚类数量，能自动识别密度较高的主题簇，同时过滤噪声文本；第四步是关键词提取，通过类内词频 - 逆文档频率(class-based TF-IDF)计算每个主题簇内的关键词权重，提取权重最高的词汇作为主题标签，例如本研究中有“电动车竞争”主题的关键词“tesla, byd, ev, sales”。

4. 研究设计

4.1. 研究步骤

研究主要包括以下三个阶段：

- (1) 数据收集与预处理：确定语料来源与时间跨度，完成数据的采集工作，运用 Python 对语料进行清洗、去重和标准化，构建语料库。
- (2) 主题发现与量化分析：根据语料特征搭建模型，运用主题建模技术自动提取语料中的核心主题分布，生成可视化结果。
- (3) 综合关联分析阶段：基于量化成果与框架理论将主题分布、地理分布、企业分布等进行质性分析，揭示其内在规律。

4.2. 语料来源

本研究选取的语料为国际主流英文媒体(包括《New York Times》《The Washington Post》《Reuters》《The Times》《Investors Business Daily》《Wall Street Journal》《The Conversation》《Toronto Star》《Azer News》《Frontier Star》等)关于中国能源企业海外投资、建设、合作及运营的报道,此类媒体文本是分析国际舆论对中国能源形象认知的重要载体[8][9]。语料收集渠道主要是 EBSCO、Pro Quest 等公开英文数据库。共收集 214 篇新闻报道,经去除重复和无效格式文本后,最终构建含 185 篇有效报道的语料库(其中包含主题-1 中的 43 篇未分类文档),聚焦能源企业在电动汽车市场、能源合作等领域的实践。

为确保语料时效性和代表性,对所收集的新闻语料进行了时间和空间分布特征的检验。基于语料元数据中的新闻发布日期统计(见表 1)可得,初始收集的新闻语料的时间跨度为 2002 年至 2025 年 7 月,其中 2020 年及以后的语料共 163 篇,占语料总数的 76.17%,具有较强的时效性和现实意义。此外,语料中所涉及地域覆盖广泛,包括欧洲、北美、南亚、非洲、中东、南美、东南亚、中亚等,能够在全球范围内代表中国能源企业的出海趋势。

Table 1. Temporal distribution of corpus
表 1. 语料时间分布统计表

年份	语料数量	占比(%)
2025	57	26.64
2024	71	33.18
2023	20	9.35
2022	10	4.67
2021	5	2.34
2020	2	0.93
2019	2	0.93
2018	1	0.47
2017	4	1.87
2016	1	0.47
2015	2	0.93
2014	0	0.00
2013	3	1.40
2012	3	1.40
2011	3	1.40
2010	6	2.80
2009	5	2.34
2008	4	1.87
2007	6	2.80
2006	0	0.00
2005	7	3.27
2004	0	0.00
2003	1	0.47
2002	1	0.47

4.3. 分析方法与参数选择

本研究采用“量化 + 质性”混合方法：在语义向量化阶段选取英文语料预训练建立 BERTopic 主题模型，以确保对英文新闻文本的语义捕捉效果，再结合 UMAP 降维技术($n_components = 5$)与 HDBSCAN 聚类算法($min_cluster_size = 3$)识别核心主题分布，最后采用类内 TF-IDF 方法提取主题关键词并生成主题 - 关键词矩阵，借助程序自动标注主题名称并结合人工核验确保命名准确性。

在 BERTopic 建模实施过程中，核心参数的设定对主题识别具有关键影响。首先，在 HDBSCAN 聚类环节， $min_cluster_size$ 被设定为 3。该参数决定了最小主题簇的规模：若取值过大，小规模但具有实际意义的新兴话题容易被忽略；若取值过小，则可能导致主题碎片化和噪声过多。通过反复对比实验，发现当取值为 3 时，能够相对兼顾主题多样性和语义连贯性。其次，在 UMAP 降维环节，本研究将 $n_components$ 设定为 5。低于该维度时，降维后的文本语义信息损失较大，导致部分主题聚类重叠；而当维度过高时，聚类结果出现不稳定现象。研究过程中的实证检验表明， $n = 5$ 能较好平衡语义保留与可视化清晰度。综合考虑实验表现与研究目标，本研究最终参数配置能够较为合理地支持主题建模的有效性、稳定性与解释能力。

5. 研究发现

5.1. 主题建模

本研究通过 BERTopic 建模识别出 15 个核心主题(含主题-1 未分类文档)，涵盖能源合作、市场竞争、环境治理等领域(见表 2)：

Table 2. Core topic distribution and keywords
表 2. 核心主题分布及关键词

主题编号	中文主题名称	英文主题名称	关键词	文档篇数
-1	未分类文档	Unclassified Documents	无	43
0	全球电动车竞争格局	Global EV Competitive Landscape	tesla, byd, ev, nio, sales, stock, vehicles, electric, auto, evs	33
1	国际油气巨头并购	International Oil & Gas M&A	oil, cnooc, gas, rosneft, oil gas, company, stake, corp, mr, unocal	19
2	中亚 - 俄罗斯能源合作	Central Asia-Russia Energy Cooperation	azerbaijan, ministry, interfax, kazakhstan, gas, said, russian, power, nuclear, en	14
3	海外高碳项目争议	Overseas High-Carbon Projects	chinas, cement, carbon, coal, countries, projects, power, overseas, carbon dioxide, dioxide	13
4	印度光伏产业政策	India's Solar Industry Policies	solar, panels, india, indias, stock, solar industry, solar cell, solar panels, money, imports	11
5	巴基斯坦电动车基建	Pakistan's EV Infrastructure	pakistan, ev, charging, ev charging, karachi, pakistani, nexus, vehicles, electric, delegation	10
6	南非技术转移合作	South Africa Technology Transfer	south africa, africa, south, transfer, vital, technology, technology transfer, minerals, renewable, supply	9
7	电池回收与储能	Battery Recycling & Storage	recycling, lfp, battery recycling, battery, energy storage, storage, hydrogen, market, applications, recovery	8

续表

8	加拿大对华 EV 关税	Canada's EV Tariffs on China	canada, canadian, evs, ev, chinese evs, tariffs, tariff, vehicles, cent, tariffs chinese	7
9	发展中国家气候政策	Climate Policies in Developing Countries	emissions, countries, trade, developing countries, 2012, climate, environmental, climate policy, developing, clean technologies	4
10	乌兹别克斯坦光伏项目	Uzbekistan Solar Projects	uzbekistan, ceec, contact, solar, modules, project, tracking, module, energy engineering, longi	4
11	关键矿物 ESG 治理	Critical Minerals ESG Governance	mining, mechanism, cobalt, mineral, metals, receives, governance, minerals, environmental social, esg	4
12	煤电与气候承诺	Coal Power & Climate Commitments	climate, coal, paris, climate change, plants, change, world, coal fired, coal plants, emissions	3
13	中国 - 巴基斯坦官方工业合作	China-Pakistan Official Industrial Cooperation	pakistan, mou, ambassador, minister, mutual, company, manufacturing plant, ahmad, federal minister, fazal	3

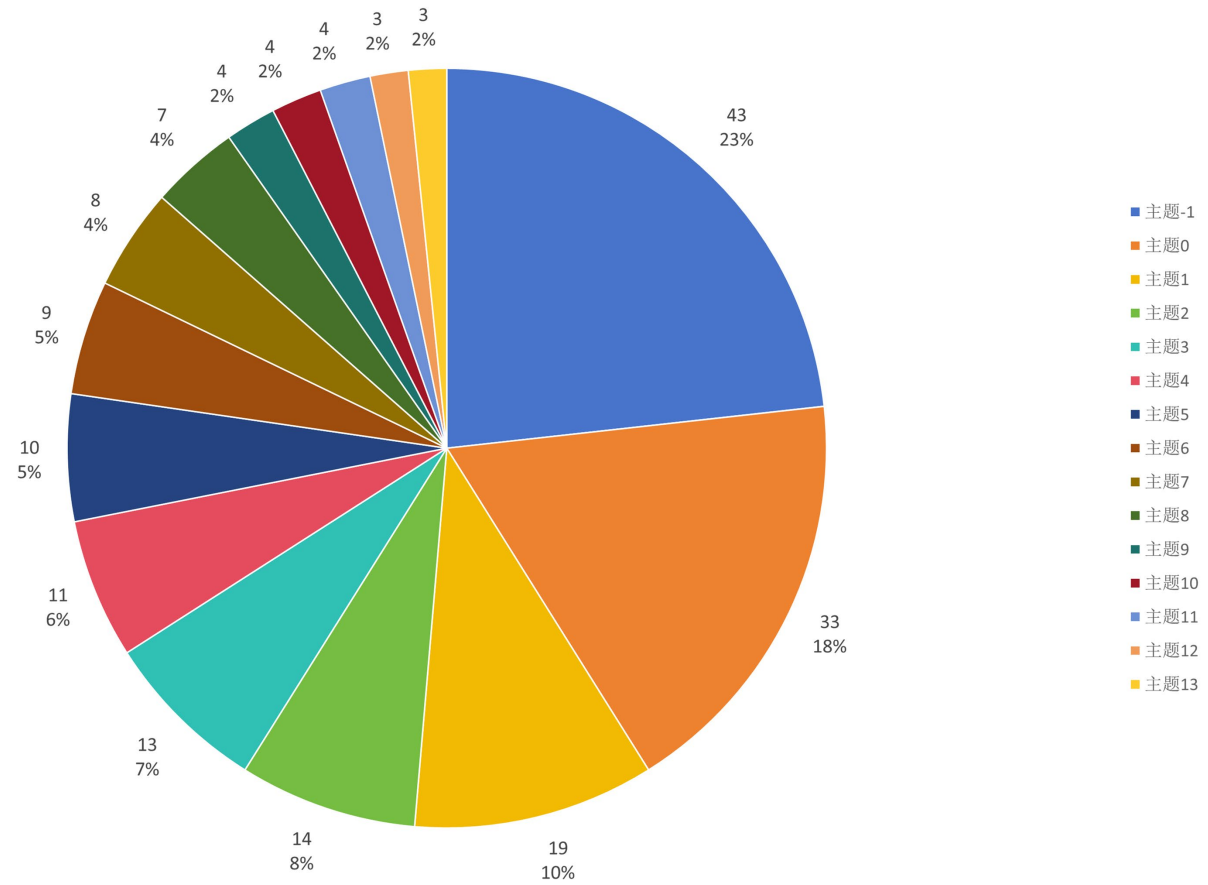


Figure 1. Topic distribution chart
图 1. 主题分布统计图

经过对 BERTopic 建模识别出的 15 个核心主题进行分析,可以清晰地观察到国际媒体对中国能源企业出海活动的关注焦点呈现出多层次、多领域的分布特征(见图 1)。本研究发现的 15 个框架,补充并系统化了先前研究的发现。Liu *et al.* (2021) [3] 曾定性指出西方媒体常将中国能源投资呈现为战略威胁,本研究则通过“加拿大对华 EV 关税”、“国际油气并购”等框架量化证实了这一倾向,并进一步揭示出“威胁”叙事已从传统地缘政治领域延伸至高科技竞争领域。同时,本研究发现了现有文献未能充分捕捉的新框架,如“关键矿物 ESG 治理”与“电池回收”,这表明国际舆论的关注点正随中国能源产业的升级而同步演化,从批判其高碳投资转向审视其在新兴绿色产业链中的角色与治理水平。这对修正学界认为国际舆论仅聚焦中国能源负面形象的片面认知具有重要价值。

这些主题共同勾勒出国际媒体对中国能源企业出海活动的认知框架,从传统能源合作到新兴技术竞争,从地缘战略布局到行业标准制定,呈现出一个多层次、多维度的复杂图景。中国企业既被视为全球能源转型的重要参与者,也在某些领域面临贸易政策和环境标准的挑战。这种多元化的主题分布不仅反映了中国能源企业出海实践的广度与深度,也为理解国际舆论态势提供了重要的分析依据。

5.2. 地理分布

本研究对中国能源企业出海英文新闻报道提及频次的地理分布情况进行统计(见表 3),发现其覆盖全球主要区域,且各区域关注焦点呈现显著差异:

Table 3. Geographic mention frequency
表 3. 地理提及频次表

地理区域	提及频次	占比(%)
欧洲	79	23.05
北美	56	16.17
南亚	41	11.87
非洲	30	8.69
中东	25	7.26
南美	23	6.65
东南亚	22	6.37
俄罗斯	22	6.37
中亚	20	5.80
澳大利亚	19	5.50
合计	343	100.00

本研究对中国能源企业出海英文新闻报道的地理分布分析揭示出明显的区域差异(见图 2),这一发现与杨志成等(2025) [10] 关于“一带一路”话语建构研究中指出的地域关注差异形成有力呼应与对话。本研究进一步深化该结论,不仅确认差异存在,更精确识别出差异的具体形态:欧美市场(欧洲与北美合计占比 39.22%)以“技术-政策博弈”为主导框架,聚焦于电动车、储能与光伏等领域的技术竞争与政策壁垒;而“一带一路”沿线地区(南亚、非洲、中东合计占比 27.82%)则更多呈现“基建-发展”框架,强调能源基础设施合作与民生改善项目。同时,俄罗斯与中亚地区(合计占比 12.17%)依托地缘邻近与资源互补,体现出传统能源与新能源双领域并行、长期战略合作关系稳定的典型特征,成为中国能源出海的稳定基本盘。这种高度分化、由地域利益驱动的框架生态表明,国际舆论并非单一整体,中国企业必须从“一体化”国际传播策略转向更具精准性和地域适应性的“差异化”传播方式。

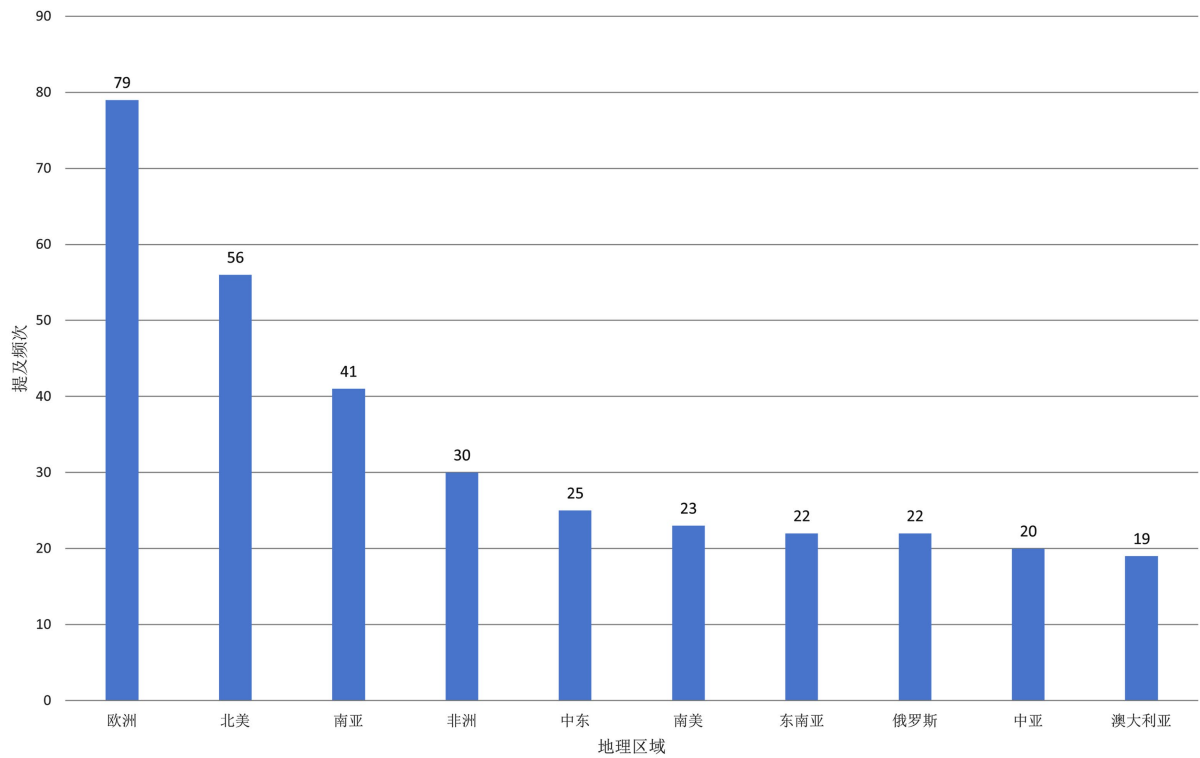


Figure 2. Geographic mention frequency chart
图 2. 地理提及频次图

5.3. 企业分布

5.3.1. 企业分布结果及其行业分类

本研究还发现对中国能源企业出海的英文新闻报道中，不同企业的提及频次存在明显差异(见表 4)，并通过绘制条形图直观地展现这种差异(见图 3)。

Table 4. Company mention frequency
表 4. 企业提及频次表

企业名称	英文名称	提及次数	占比(%)
比亚迪	BYD	35	33.98
中国石油天然气集团有限公司	CNPC	13	12.62
中国石化	Sinopec	12	11.65
宁德时代	CATL	10	9.71
大唐集团	Datang	7	6.80
中国海洋石油集团有限公司	CNOOC	6	5.83
中国能源建设集团有限公司	CEEC	6	5.83
隆基绿能	LONGi	5	4.85
中国电力建设集团有限公司	Power China	5	4.85
国家电网	State Grid	4	3.88
国家电力投资集团有限公司	SPIC	3	2.91

续表

天合光能	Trina	3	2.91
华能集团	Huaneng	2	1.94
中国有色金属矿业集团	CNMC	2	1.94
总计	-	103	100.00

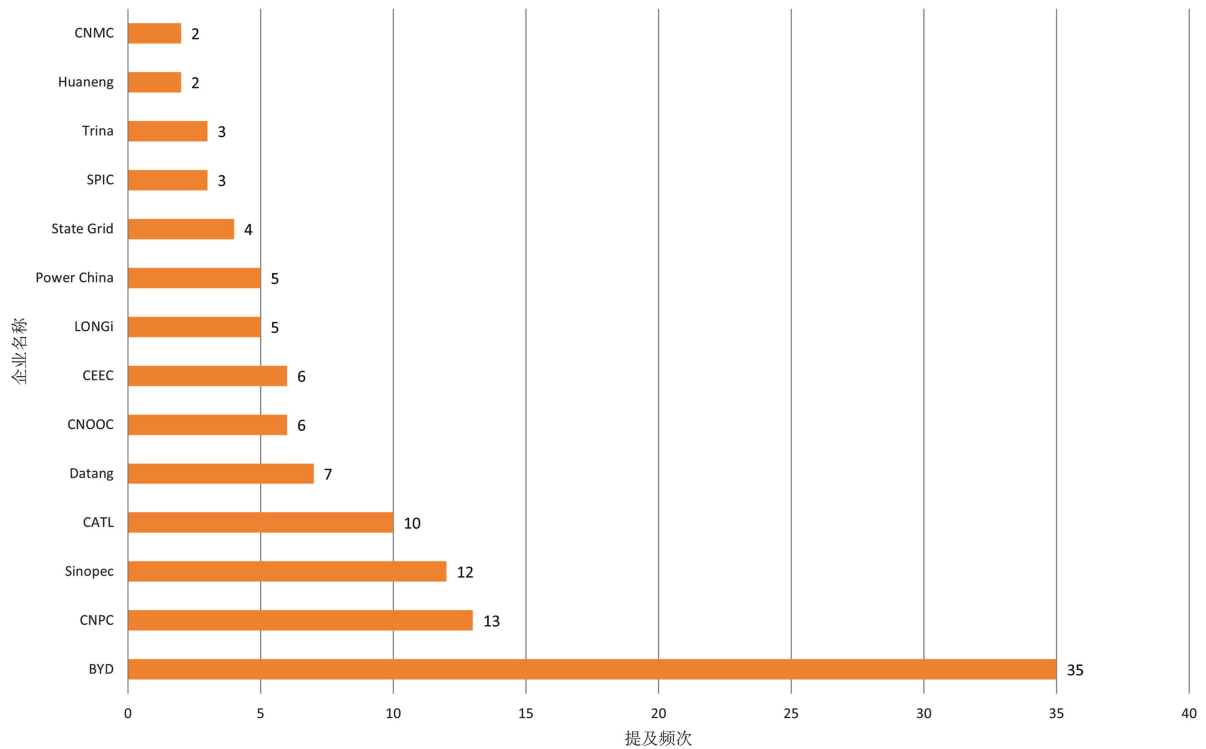


Figure 3. Company mention frequency chart
图 3. 企业提及频次图

此外，本研究将企业按主营业务分类，能更清晰地看出媒体关注的行业分布(见表 5):

Table 5. Industry sector distribution
表 5. 行业领域分布表

行业领域	代表企业	提及总次数	合计占比
新能源汽车	比亚迪、宁德时代	45	43.69%
传统油气	中石油、中石化、中海油	31	30.10%
电力建设与服务	中国电建、中国能建、国家电网	15	14.56%
光伏产业	隆基、天合、国家电投	11	10.68%
传统发电	大唐、华能	9	8.74%
矿产资源	中国有色	2	1.94%

特别需要声明的是，我们在企业分析过程中发现通用词组“China Energy”在不同语境下具有高度歧义性，它可能指代中国能源行业、国家能源政策，或在非正式场合下作为国家能源投资集团(CEIC)的简

称, 以及作为某些能源企业如中国能建(CEEC)全称中的一部分, 不应简单归类为国能集团的名称。为确保企业实体出现频次统计的准确性与可靠性, 避免数据污染, 本研究采取了保守的数据清洗策略: 在分析中仅将那些使用企业官方注册名称(如 CEEC)的实例计入相应企业的频次统计, 而将所有孤立出现的、歧义的“China Energy”实例视为噪声数据并从企业分析中排除。

5.3.2. 企业梯队分析

根据提及频次, 可以将以上企业分为三个明显的梯队:

第一梯队(绝对焦点): 比亚迪(BYD, 33.98%): 独占超三分之一的频次, 是国际舆论的关注核心, 表明中国在电动汽车领域的突破和全球竞争(尤其是与特斯拉的竞争)是当前最吸引国际媒体目光的题材。

第二梯队(重要参与者): (1) 中国石油(CNPC, 12.62%)、中国石化(Sinopec, 11.65%): 作为中国传统的能源巨头, 其海外业务、合作与并购依然是稳定的新闻源, 但关注度已明显不如新能源明星企业; (2) 宁德时代(CATL, 9.71%): 作为全球动力电池领域的绝对龙头, 其技术突破、国际工厂布局和客户合作备受关注[11], 是比亚迪之外的另一新能源亮点。

第三梯队(细分领域代表): (1) 传统发电与工程企业: 大唐、国家电网、中国能建、中国电建、华能等; (2) 光伏与新能源巨头: 隆基绿能、天合光能等; (3) 专业领域企业: 中海油、中国有色金属矿业集团等。

首先, 媒体的关注点覆盖了从上游的矿产资源、到中游的电池制造、光伏组件, 再到下游的整车制造和能源基建, 这表明中国能源出海已经演化成一个“全产业链”形态, 而不再仅仅是过去单一产品的出口。更重要的是, 企业关注度的梯队分布, 是对 Zhang & Andrews-Speed (2020) [1]所描述的中国能源海外投资“传统模式”的实质性修正。其研究仍以传统油气巨头为核心, 而本研究的数据清晰地表明, 新能源汽车和光伏的提及占比高达 54.37%, 已经显著超过传统油气(30.10%), 这说明国际媒体的框架焦点已发生历史性转移: 新能源科技企业如比亚迪、宁德时代等逐渐成为框架体系中的相对核心。这是一个标志性的信号, 反映了中国在全球能源叙事中的角色, 正被国际媒体从“资源购买者”和“工程承包商”重新框架化为“技术挑战者”与“市场规则重塑者”。这一转变是根本性的, 它要求学术研究与企业实践都必须将新能源领域的国际传播置于更优先的地位。

6. 小结

本研究通过 BERTopic 主题建模技术对国际英文媒体关于中国能源企业出海的报道进行了系统分析, 研究结论不仅回应了引言中提出的三个核心问题, 更与现有文献形成了深入的对话与拓展, 体现出多方面的学术贡献。首先, 在主题识别层面, 本研究识别出 15 个核心主题, 涵盖从传统油气合作到电池回收、关键矿物 ESG 治理等新兴领域。这一发现补充并拓展了 Liu *et al.* (2021) [3]关于西方媒体将中国海外能源投资框架化为“战略威胁”的结论。本研究不仅证实了这一框架在传统领域的延续性, 更重要的是揭示了国际舆论框架的演化趋势: 即关注点已从单一地缘政治风险, 延伸至高科技竞争与绿色治理等新领域。这表明, 随着中国能源产业升级, 国际媒体的叙事逻辑也在同步调整, 修正了以往研究认为国际舆论仅聚焦负面形象的片面认知。其次, 在地理分布层面, 本研究验证了杨志成等(2025) [10]关于“一带一路”话语建构中存在地域分异的观点, 但进一步深化了这一结论。本研究通过精确的量化分析揭示出差异的内在结构: 欧美市场以“技术-政策博弈”框架为主导, 而“一带一路”沿线国家则以“基建-发展”为叙事主轴。这种由地域利益驱动的高度分化框架生态, 完善了将国际舆论视为同质整体的简化观点, 建立更精细化的区域舆论分析框架。最后, 在企业关注维度上, 本研究发现国际媒体报道焦点发生历史性转变, 新能源企业的媒体占比(43.69%)已显著超过传统油气企业(30.10%)。这一发现对 Zhang & Andrews-Speed (2020) [1]所描述的中国海外能源投资“传统模式”形成了实质性修正。研究表明, 中国在

全球能源叙事中的角色正从“资源购买者”和“工程承包商”被重新框架化为“技术输出者”与“市场规则挑战者”。这一转变不仅呼应了全球能源转型趋势，也挑战了以往研究中以传统能源为重心的分析范式，凸显出新能源领域国际传播机制研究的紧迫性与重要性。

本研究的结论对中国能源企业的国际传播与战略布局具有重要实践意义。中国能源企业应建立区域适配的差异化传播策略，在欧美市场聚焦技术领先与合规叙事，在“一带一路”国家突出与当地的民生发展与建设合作，在地缘邻近区域巩固稳定互信的基调。同时，企业应提升在绿色低碳议题上的话语能见度，塑造负责任的清洁能源技术领导者形象，并持续优化企业形象建构中对自身的角色定位，将企业出海实践与可持续发展目标相衔接，构建更具包容性的国际话语体系。

不过，本研究仍存在一些局限性。在语料数据方面，样本主要来自主流英文媒体，未广泛纳入社交媒体等多源文本，且所收集样本量有限，难免影响结论的全面性与可信度。在分析方法上，未能深入考察报道的情感倾向和立场差异，也缺乏对文本话语策略的微观解析。未来研究可进一步扩大语料范围和规模，优化语料收集方法，融入多类型、多语种文本；同时结合情感分析量化媒体态度，加强微观文本分析，从而更全面揭示国际舆论的建构机制与影响。

综上所述，本研究通过融合框架理论与深度学习文本分析方法，实现了对国际媒体有关中国能源企业出海报道的大规模、结构化解析，不仅在实证层面补充和修正了现有文献的观点，更在理论层面丰富优化了过于简单或滞后的学术认知。研究成果为构建更精准、更具解释力的中国能源企业国际形象分析框架提供了重要依据，也为后续学术研究与企业出海实践提供了一定参考。

基金项目

中国矿业大学(北京)大学生创新训练项目资助(项目编号: 202508001)、中央高校基本科研业务费资助。

参考文献

- [1] Zhang, L. and Andrews-Speed, P. (2020) China's Global Energy Strategy: Policy, Practice and Prospects. *Energy Policy*, **145**, Article ID: 111792.
- [2] Freitas, I.B., Jacob, J., Wang, L. and Li, Z. (2022) Energy Use and Exporting: An Analysis of Chinese Firms. *Journal of Evolutionary Economics*, **33**, 179-207. <https://doi.org/10.1007/s00191-022-00796-1>
- [3] Liu, X., et al. (2021) Western Media Framing of China's Overseas Energy Investments. *Energy Research & Social Science*, **78**, Article ID: 102142.
- [4] Goffman, E. (1974) *Frame Analysis: An Essay on the Organization of Experience*. Harper & Row.
- [5] Entman, R.M. (1993) Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, **43**, 51-58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- [6] Grootendorst, M. (2022) BerTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure.
- [7] Okazaki, H. and Takahashi, H. (2025) Analysis of Extracting Corporate Trends from News Data Using BERTopic. *New Generation Computing*, **43**, 1-16. <https://doi.org/10.1007/s00354-025-00290-9>
- [8] 董其鑫. 中国能源形象的他建与自塑: 澳大利亚和中国主流媒体报道比较[D]: [硕士学位论文]. 海口: 海南大学, 2022.
- [9] 张越. 批评话语分析视域下中俄媒体的中国能源形象研究[D]: [博士学位论文]. 海口: 海南大学, 2024.
- [10] 杨志成, 宋周莺, 冯靖翔, 等. 西方舆论对绿色“一带一路”的话语建构[J]. *地理研究*, 2025, 44(1): 1-20.
- [11] 汪真祺. “宁德时代”国际化经营策略研究[D]: [硕士学位论文]. 南宁: 广西大学, 2024.