

# India's long shadow of coal: Media framing of coal amidst transition calls

Arne Arens<sup>a</sup>, Charlotte Sophia Bez<sup>b,c</sup> <sup>\*</sup>

<sup>a</sup> Hertie School, Friedrichstraße 180, Berlin, 10117, Germany

<sup>b</sup> Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, Telegrafenberg A 31, Potsdam, 14412, Germany

<sup>c</sup> The New School, 66 West 12th Street, NY, 10011, USA

## ARTICLE INFO

Handling Editor: Henrik Lund

Dataset link: <https://doi.org/10.5281/zenodo.18348399>

### Keywords:

India  
Coal  
Energy transition  
Discourse analysis  
Text-as-data  
Topic modelling  
Sentiment analysis  
Named entity recognition

## ABSTRACT

India's coal sector plays a central role in debates on energy transitions, yet systematic evidence on how media frame this transition remains limited. We address this gap by analysing ten years of English-language news coverage using a combined pipeline of dynamic topic modelling, sentiment analysis, and named entity recognition. Across more than 6 thousand articles, we map the evolution of topics, tones, and actor prominence. Environment-related narratives account for only about 13% of coverage, and governance and social issues around 12%, indicating their secondary position in the broader discourse. The energy transition appears largely marginal, typically embedded within frames of coal supply and energy security. Attention spikes around environmental or governance events are short-lived and do not shift narrative weight. Positive tones cluster around topics tied to state and corporate actors emphasising continuity and expansion. We show the coal complex's discursive dominance and find China to be the most prominent external actor in transition-related coverage. The results provide an evidence-based foundation for subsequent analyses of policy windows and framing strategies around India's coal phase-down and clean-energy transition.

## 1. Introduction

India's energy system remains profoundly coal-dependent. In 2022–23, coal and lignite together generated nearly 77% of the country's electricity.<sup>1</sup> Beyond power generation, coal underpins a wide range of energy-intensive industries such as steel, cement, and bricks, and continues to be viewed domestically as essential for economic development and energy security. If coal combustion persists at current levels, there cannot be a meaningful energy transition. Yet, phasing out coal in India is deeply contested and political [1], shaped by contested narratives that reflect divergent values and interests. National discourse commonly frames coal in terms of energy crisis and resource security [2]. Global climate finance actors approach coal primarily through the lens of transition imperatives, often overlooking regional socio-economic dependencies. These conflicting frames highlight the need to understand how policy options are articulated and legitimised in public debate.

This international context has become increasingly salient, as India's domestic coal politics now intersect directly with global climate

finance negotiations and expectations around energy transition pathways. Worldwide research indicates that green finance can play a significant role in reducing coal dependence while accelerating renewable energy deployment [3]. Nevertheless, this financial dynamism has not translated into a retreat from coal. India continues to expand coal extraction and generation capacity, underscoring the coexistence of renewable expansion and coal entrenchment. The failure of the proposed Just Energy Transition Partnership (JETP) between India and donor countries illustrates this tension. Despite the G7's 2022 commitment to support India's decarbonisation efforts, negotiations stalled and the partnership was terminated at the end of 2024.<sup>2</sup> As a result, India remains deeply invested in coal even as renewable energy capacity has expanded rapidly, reaching nearly 210 GW in 2024 [4]. It is now second only to China in planned coal capacity additions, with almost 100 GW under construction or in pre-construction [5].

Understanding these dynamics requires situating India's coal sector within its broader political economy. Coal supports the livelihoods of an estimated 13–15 million people and is foundational to the economies of at least five states [6,7]. This dependency is institutionalised through

\* Corresponding author at: Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, Telegrafenberg A 31, Potsdam, 14412, Germany.

E-mail address: [charlott@pik-potsdam.de](mailto:charlott@pik-potsdam.de) (C.S. Bez).

<sup>1</sup> <https://energy.economictimes.indiatimes.com/news/renewable/india-still-heavily-dependent-on-coal-takes-77-per-cent-share-in-total-energy-generation-in-fy-23-says-mospi/108644839> (accessed 01/12/2025).

<sup>2</sup> [www.cleanenergywire.org/news/india-donor-countries-give-just-energy-transition-partnership-german-official](http://www.cleanenergywire.org/news/india-donor-countries-give-just-energy-transition-partnership-german-official) (accessed 27/04/2025).

Coal India Limited (CIL), the world's largest coal producer, and a state-led coal economy rooted in the 1973 nationalisation, which prioritised public-sector extraction for energy security [8]. Since the 1990s, a neoliberal coal economy has emerged alongside it, enabling private participation and the rise of conglomerates such as Adani and Tata [9,10]. This shift has generated what Kumar [11] terms "fossil neoliberalism" and contributed to a "new coal geography" characterised by coastal coal-fired plants, expanded import reliance, and reconfigured trade flows [12].

Coal-rich states such as West Bengal, Jharkhand, Chhattisgarh, Odisha, and Madhya Pradesh are further shaped by extensive informal labour, precarious working conditions, and extractive practices marked by state violence and land dispossession [13]. These realities complicate the application of just transition frameworks developed largely in the Global North. Because the burden of transition is regionally concentrated, the socio-economic disruption of a coal decline would fall disproportionately on specific districts [14]. As Chandra et al. [7] argue, the faltering of India's JETP negotiations reflects a deeper disconnect between global coal phase-down expectations and the domestic reality of a growing coal industry.

Against this backdrop, research on traditional energy sources such as coal remains indispensable, as fossil fuels continue to shape the political and economic constraints within which transitions must occur [15]. At the same time, India has explored alternative fuel strategies (e.g., biofuels, hydrogen, and electric mobility) to reduce transport emissions and strengthen energy security, though these pathways face considerable technological and economic constraints [16]. For countries in the Global South, where coal use is expanding to meet energy demand, transitions must therefore prioritise affordability, accessibility, and fairness (see also [17]).

Yet material constraints alone do not determine transition feasibility. Public narratives shape how problems are defined, who is responsible, and which policy options appear viable. Indian media is a key arena where competing interpretations of coal, energy security, and transition futures are formed and contested. Understanding this discursive landscape is essential for assessing the political feasibility of climate-policy and climate-finance interventions, and for identifying entry points where such proposals may gain traction. Hence, research on traditional energy sources such as coal remains indispensable, as these fuels continue to underpin energy systems and shape the political and economic constraints within which transitions unfold [15].

Achieving an energy transition in India that leaves nobody behind is a difficult task across multiple dimensions. Practically, India would need to dramatically ramp up its renewable capacity additions to a globally unprecedented level, taking into account the levels of growth forecasted by the Indian government. India's highly regional coal production adds an important social dimension to the complexity of the task, as creating alternative income streams for millions would be immensely challenging and require long-term industrial and labour strategies. For countries in the Global South, where coal use is still expanding to fuel economic growth and growing energy demands, energy transitions must prioritise affordability, accessibility, and fairness (see also the energy justice index developed by Aperi et al. [17]).

Hence, it becomes critical to understand how these priorities are articulated, contested, or sidelined in public debate. Transitions remain complex not only due to material constraints but also due to struggles over meaning and legitimacy. At its core, this study is motivated by the theoretical understanding that energy transitions are shaped not only by material constraints but also by policy narratives that frame problems, attribute causes, and legitimise solutions [18,19]. News media actively structure these narratives by functioning as agenda-setting arenas for elites [20,21] and amplifying particular storylines around risk, opportunity, and justice [22]. Such narrative dynamics are politically consequential: they shape public perceptions, coordinate incumbent interests, and delimit the range of feasible policy options [23–25].

Mapping discourses is thus essential for identifying entry points for climate policy across mitigation, adaptation (see, e.g., [26]), and loss and damage (see, e.g., [27]). Building on the view that diversity in discourse guides research and policy priorities [26], we examine a decade of Indian English-language news to analyse the coal landscape and to see where, how, and by whom energy transition considerations are articulated within in. In doing so, we address a key research gap because transition-related narratives often surface only at the margins of wider reporting on coal and energy security, making it difficult to assess their political resonance or durability. It is not well understood whether environmental concerns receive episodic rather than sustained media attention. Moreover, existing studies do not examine how the structural dominance of coal-related actors shapes the narrative architecture of media discourse, thereby conditioning which transition pathways appear politically plausible. Prior work has examined India's transition pathways and their socioeconomic implications [28,29], industrial decarbonisation [30], and system-level mitigation constraints [31], and isolated studies have explored media perceptions in other national contexts [32], no research has provided a longitudinal analysis of how India's coal (transition) is framed in national media. Our study addresses this gap. Our sequential pipeline links topics to tone and actors, which allows us to observe agenda setting and discursive power in practice: which narratives secure sustained attention, which actors can credibly perform urgency or stability, and which policy options become thinkable. The contribution is twofold. First, we provide a systematic map of long-run narrative dynamics that identifies how major topics emerge and evolve. Second, we identify the attitudinal framings and actor groups that condition the reception of transition proposals. Together, they offer practical guidance on where climate-policy dialogue and targeted climate-finance mechanisms can be most effectively anchored.

We depart from several hypotheses. Based on existent literature, we expect narrative continuity, i.e., transition debates in India to be typically anchored in development and energy-security logics, rather than standing as an autonomous agenda. As a new, empirical hypothesis, we pose that positive tones cluster around business topics, which includes a business-led narrative of the energy transition, while purely environmental frames are more negative. This reflects that market-oriented stories tend to emphasise opportunity, while environmental reporting more often highlights risk, conflict, or loss. Ex ante, we do not know whether event-driven spikes in governance and pollution coverage produce lasting shifts in narrative, e.g., though policy changes. It is hence unclear how short-lived media attention cycles translate into lasting policy effects. Lastly, we hypothesise that the dominant actors are the incumbent coal industry stakeholders, as India's energy sector is institutionally centralised and heavily state-industry driven. Non-state actors who offer counter-discourses, thus remain secondary, in line with evidence of limited agenda-setting power among civil-society organisations in the coal domain. To answer these questions, we propose an innovative sequence of automated content analysis, especially suitable for energy-transition discourse where meaning depends on both *what* is being discussed and *who* is speaking with *what* tone.

This paper is structured as follows. Section 2 situates our study in the broader literature on coal in India, energy policy narratives and text-as-data applications. Section 3 describes the LexisNexis data and carefully lays out the different methods involved. Sections 4–6 present the results on topics, tones and actors. Section 7 discusses our results against the background of policy narratives conceptualisation and lessons learned for climate commitment frames such as the JETP. We also acknowledge several limitations of our study. Finally, Section 8 concludes.

## 2. Literature

This paper builds on energy policy narratives and text-as-data applications. In the following, we carefully lay out each of them.

## 2.1. Energy policy narratives

Policy narratives are central to policy change because they frame problems, attribute causes, and legitimate solutions, thereby shaping public opinion and coordinating the behaviour of incumbent actors [18, 19]. In the context of energy transitions, such narratives play a critical role in shaping how incumbent energy systems, emerging alternatives, and transition trade-offs are publicly understood. News media are a key arena in which these dynamics unfold, as they act as a conduit through which sustainability narratives are amplified, contested, or reframed. Understanding India's media discourse is therefore essential for assessing the political feasibility of a national energy transition and its broader implications for environmental governance. At the same time, major economies are undergoing significant shifts in their energy paradigms, making it crucial to understand how established narratives evolve in response to emerging transition pressures [15].

A growing literature shows that media do not merely report these narratives but actively structure policy discourse, functioning as agenda-setting arenas for elites [20,21] and amplifying particular storylines [22]. This dynamic is especially salient for energy transitions away from coal, where competing accounts of risk, opportunity, and justice vie for dominance. Media outlets provide venues in which advocates stage discursive performances, signalling credibility and urgency, mobilising coalitions, and marginalising alternatives. The results is a coherent yet often polarised frames around climate-related issues [23–25]. Narratives thus do not simply reflect policy priorities; they help constitute them, shaping both public perceptions and the range of politically feasible options. Further, so-called discursive power (see [33], for a definition) operates by setting the terms of debate, and whether and how it changes over time. Such power privileges certain problem definitions, actors and rules of evidence, meaning that some policy options appear sensible while others become unthinkable. In India, energy-transition talk would need to unfold within the broader coal narrative. The landscape of security and economic growth logics that subordinate decarbonisation aims and narrow the range of plausible alternatives must thus be understood better. We aim to uncover such a landscape for India to ultimately reveal the conditions under which political support or resistance to a just energy transition is articulated, thereby complementing contributions such as Ordonez et al. [34], Blankenship et al. [1] and Shukla and Swarnakar [25], who focus on employment dynamics and justice concepts, through an analytically distinct focus.

## 2.2. Text-as-data applications

Energy research increasingly recognises the value of text-as-data methods for tracking large-scale trends and analysing narrative framing in energy and climate debates [15,35,36]. As energy strategies evolve, these shifts are increasingly reflected in large textual datasets, including policy documents, parliamentary debates, and media coverage. Park [15] argue that text-analytic approaches are particularly well suited to capturing emerging trends in energy policy and public discourse that are difficult to observe through conventional qualitative methods alone.

Within this growing body of work, machine learning-based text analytics have been highlighted as a promising but still underdeveloped area of policy research. A recent systematic review by Kuenzler et al. [37] underscores the potential of text-as-data approaches for uncovering policy narratives and advocates for the further development of new methodologies. The authors stress that large-scale text analytics remain in their infancy, offering substantial untapped potential for future research (p.14). Among the available methods, topic modelling has received particular attention, as it enables researchers to identify shifts in thematic focus and examine how policy issues are reframed or prioritised across contexts and over time. Several studies demonstrate how topic modelling can generate actionable insights for energy policy analysis. For example, Liu [35] apply BERTopic to identify thematic

gaps in existing energy policies. Such approaches are particularly effective when applied to large textual corpora, allowing for systematic analysis at scale while maintaining interpretability [38–40].

In parallel, discourse analysis has been widely used to examine how narratives evolve, compete, and shape public opinion and policy outcomes in the energy sector [32,41]. Jiang and Raza [41] demonstrate that analysing the discursive functions of policy texts can reveal whether policies perform their intended roles in guiding energy transitions. More recent work has integrated discourse analysis with text-as-data methods to examine coal transitions in other national contexts. For instance, Bez et al. [42] show a systematic map of coal discourse in South Africa. These studies highlight how dominant framings (such as portraying coal as a cornerstone of economic stability) can directly influence political will and the design of financial mechanisms in international climate politics, including the JETP.

Building on this literature, our framework advances a scalable and interpretable text-as-data pipeline that integrates three core components: topics, tones, and actors. Topic modelling is used to identify how narratives emerge, persist, or are replaced over time. Sentiment analysis captures the emotional framing associated with each topic, offering insights into how narratives resonate with different audiences or align with particular policy objectives [22]. Polarised sentiment can further signal moments of tension or contestation in policy debates. Named Entity Recognition identifies the actors associated with specific narratives and clarifies their roles within the discourse. Finally, computational results are qualitatively validated and interpreted, combining large-scale empirical analysis with contextual nuance.<sup>3</sup>

## 3. Data and methods

We first describe our data, detailing the source of the news media articles and the preprocessing steps (in 3.1). We then present the methodological pipeline consisting of several computational text-as-data tools, combined with qualitative methods (in 3.2). Each result section leverages a different set of computational tools.

### 3.1. Data

The initial dataset comprises 127,945 Indian English-language newspaper articles, sourced from the LexisNexis API in December 2023. LexisNexis is a leading global provider offering extensive news databases, which are widely used by researchers, journalists, and professionals. Articles were retrieved using a keyword search for 'coal' across headlines in Indian publications, from 2013 onwards.

The dataset was refined in several stages, as visualised in the Sankey graph (see Fig. 1). First, articles misclassified in the LexisNexis database, such as those originating from other countries or consisting of parliamentary records, were removed. Publications were then filtered based on readership data from the Indian Readership Survey (IRS), ensuring representation of the most widely read English-language newspapers in India (see Table A.1). From the top ten newspapers, we retained seven (resulting in  $n = 12,809$ , see Fig. 1), including The Economic Times, a business daily consistently ranked among the top ten. (Consult Fig. A.3 in Appendix A for the new outlet distribution in our dataset.) All news outlets are published daily, as shown in column 2 in Table 1. Several notable newspapers (Mumbai Mirror, Mid Day, The Tribune, and Deccan Chronicle) were absent from the LexisNexis archive during our study period, a gap we identified by comparing our corpus against national readership rankings. Deccan Herald was additionally excluded due to its minimal representation in the archive.

<sup>3</sup> This approach aligns with [43], who combine topic modelling with qualitative narrative coding, but extends the methodological scope to provide a more comprehensive analytical toolkit.

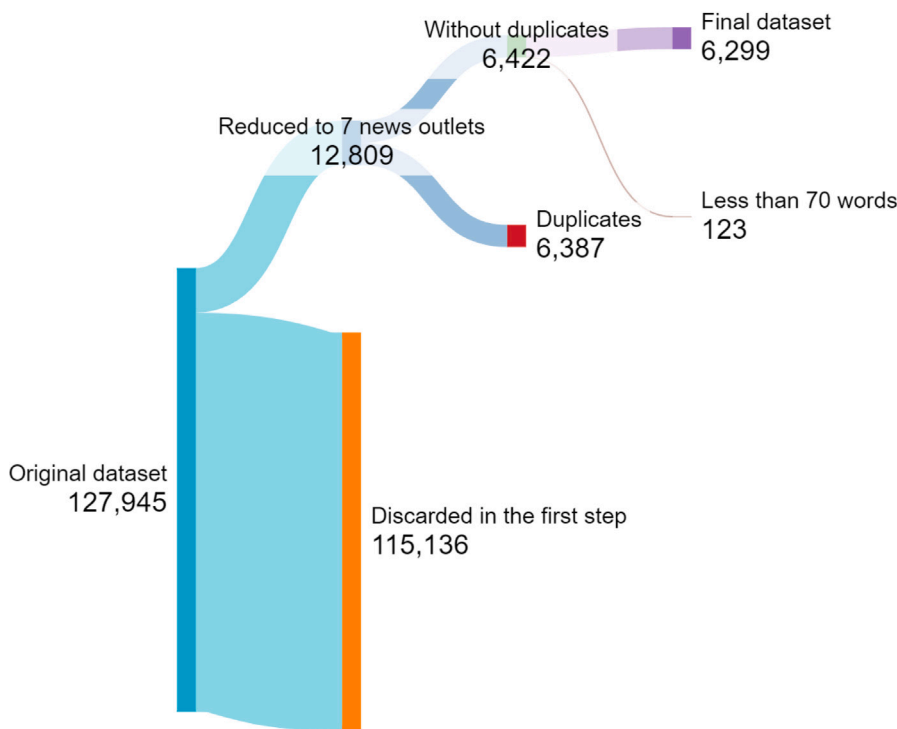


Fig. 1. Sankey graph illustrating the sequential reduction of the dataset through filtering for 7 most relevant news outlets, duplicate and false positive check and removal of very short articles. The final dataset consists of 6,299 articles.

To prepare the resulting text corpus for analysis, we first removed artifacts such as web links, embedded HTML, and non-alphabetic characters. We then conducted a duplicate analysis using TF-IDF vectorisation and pairwise cosine similarity, classified article pairs with similarity scores above 0.8 as duplicates and retained only the earliest occurrence. Articles with missing dates or fewer than 70 words were also excluded. The remaining texts were lowercased and tokenised using a regular-expression tokenizer that retained alphabetic tokens only. Stop words were removed using the NLTK English list augmented with additional high-frequency terms, calendar references (days, months, years), and domain-specific vocabulary. Tokens containing digits were removed and the Porter stemmer was applied to standardise word forms, after which additional high-frequency residual stems were removed. All preprocessing was conducted in Python (version 3.11) using standard libraries including scikit-learn, NLTK, and pandas. Following these preprocessing steps, the dataset comprised 6,299 documents (shown in purple in Fig. 1).

To better understand our news corpus, we also accessed political orientations of the included publications, following the classifications of Sen et al. [44]. The Times of India and The New Indian Express were identified as Bharatiya Janata Party (BJP)-aligned, while The Hindu and The Indian Express leaned towards the Indian National Congress (INC). Hindustan Times was considered neutral but anti-INC, and The Telegraph was classified as anti-BJP and anti-INC. These findings were further corroborated by Media Bias/Fact Check (MBFC) ratings [45], which generally categorise the included outlets as left-centre biased, except for The Times of India and The Economic Times, both rated right-centre. None of the included newspapers were rated at either extreme of the bias spectrum (see the third column in Table 1).

In Appendix A, we also show the distribution of article length (Fig. A.2). Based on our final dataset, Fig. 2 illustrates the monthly number of articles published from early 2013 to late 2023. Publication volume peaked sharply in early 2021, with over 180 articles in a single month. Additional periods of heightened activity occurred in 2015 and 2022,

Table 1

List of news outlets included in the dataset, type and MBFC bias ratings. Source: India Readership Survey over the years 2013, 2014, 2017, and 2019, Media Bias/Fact Check.

News outlet	Type	MBFC bias rating
The Times of India	Daily newspaper	Right-centre
Hindustan Times	Daily newspaper	Left-centre
The Hindu	Daily newspaper	Left-centre
The Telegraph	Daily newspaper	N/A
The Indian Express	Daily newspaper	Left-centre
The New Indian Express	Daily newspaper	Left-centre
The Economic Times	Business daily newspaper	Right-centre

while noticeable declines were observed around 2017–2019. Overall, the data reveals substantial variability in article frequency, suggesting shifts in media attention or content production trends over time. This time span over a decade was selected both for methodological coherence and because it spans several major developments in India’s coal and climate policy landscape, including the period leading up to and following the Paris Agreement and the emergence of JETP negotiations.

The data have two possible limitations. First, we include all Indian news articles that mention “coal” without ex-ante thematic filtering. This maximises longitudinal coverage of coal discourse but admits false positives. We mitigate this with topic modelling and targeted qualitative checks. Second, the corpus is limited to English-language newspapers. This necessarily omits many subnational and vernacular perspectives, especially from coalfield regions, but it is consistent with our research focus on national policy debate, where English is the prevailing medium in government communication, business journalism, and elite public spheres. The language restriction is thus a deliberate choice to analyse agenda setting and policy framing, while compromising broader representativeness across India’s linguistic landscape. Finally, the dataset ends in late 2023 and does not capture media coverage from 2024, including the final collapse of India’s JETP negotiations.

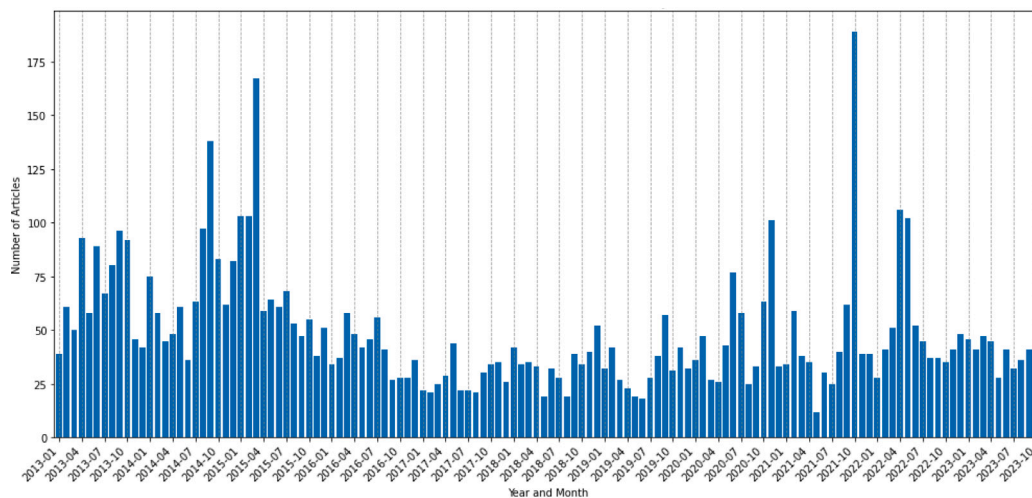


Fig. 2. Distribution of articles per month, from early 2013 to late 2023.

This reflects the timing of our data collection and analysis, but it limits our ability to examine narrative responses to this key event.

### 3.2. Methods

Text-as-data approaches are vital for uncovering policy narratives, yet many studies rely on a single method, which narrows the analytical scope. We argue that combining methods sequentially strengthens automated content analysis, especially for energy-transition discourse where meaning depends on both *what* is being discussed and *who* is speaking with *what* tone. We therefore estimate topics first using dynamic topic modelling (DTM), a dynamic specification to track trajectories over time, and then apply state-of-art sentiment analysis and named entity recognition (NER) to those topic. This sequencing is deliberate. Deriving sentiment and actors conditional on topics yields interpretable, topic-level measures rather than corpus-wide aggregates. It reduces noise from semantic ambiguity and allows us to link differences in tone and key actors with thematic differences. Fig. 3 illustrates our methods pipeline: DTM supplies the topics (in green, presented in Section 4); sentiment (in blue) and NER (in red) operate on them. Their intersections capture combined outputs, i.e. “Tones across topics” (Section 5) and “Actors across topics” (Section 6). We complement these computed measures with targeted qualitative reading of articles to validate model outputs and refine interpretation.

In what follows, we outline each method in a separate subsection and describe how it is applied to the corpus.

#### 3.2.1. Dynamic topic modelling

We employ dynamic topic modelling (DTM) to explore the temporal evolution of topics. The idea is to quantitatively associate documents with various topics through assigned weights (topic scores), thus transforming the raw textual data into structured, thematic patterns [39]. DTM utilises unsupervised machine learning to detect recurring themes within large datasets of text. Based on the assumption that frequently co-occurring words likely share semantic relationships, topic models abstract and simplify the complex, high-dimensional data of word occurrences in documents into two interpretable forms [46]. The first form is a distribution of terms that define each topic, while the second form reflects how these topics are distributed across the documents themselves. By analysing word frequencies and co-occurrences, topic models induce clusters-topics, each characterised by a distinctive term distribution.

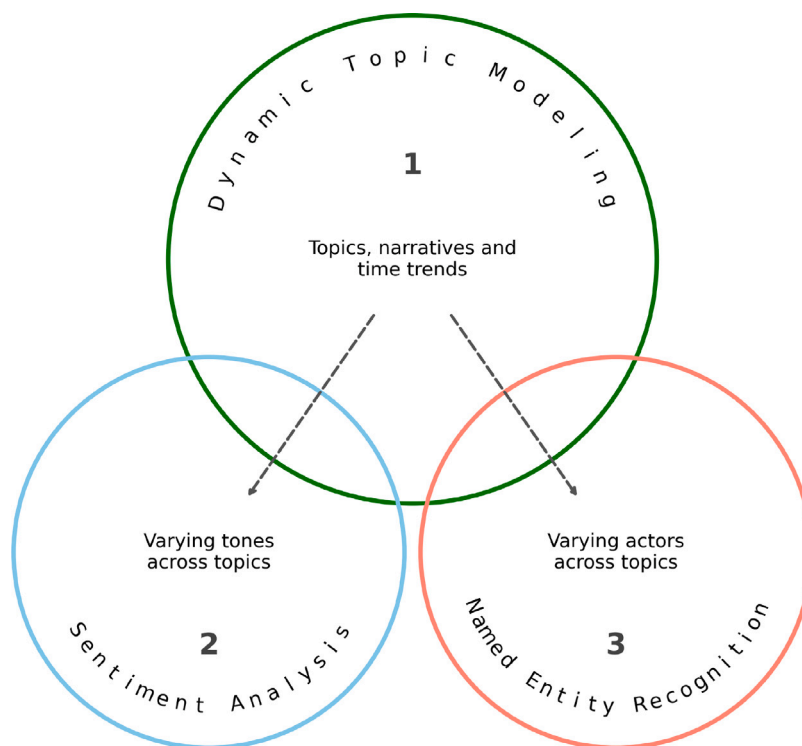
DTM further provides a robust framework for tracing how discussions about coal have evolved over time. Unlike Latent Dirichlet Allocation (LDA), which treats topics as static, DTM incorporates a

temporal dimension, capturing shifts in language, discourse, and emphasis across periods [46]. This method is particularly suited to our analysis of the Indian discourse around coal, as it allows us to examine how environmental, economic, and policy-related discussions around coal have developed alongside national and global events, up until the official failure of the JETP. DTM quantitatively describes these shifts by adjusting the prevalence and salience of topics based on the temporal metadata associated with each document. This method therefore offers a deeper and more granular understanding of discourse dynamics than traditional static topic modelling methods [47].

A crucial step in topic modelling is selecting the right hyperparameters. For DTM, the most important is the number of topics,  $k$ . Determining an appropriate  $k$  always involves qualitative analysis, as noted by Müller-Hansen et al. [39] and Törnberg and Törnberg [48]. We first evaluated model quality by computing Coherence and Exclusivity scores for different values of  $k$  sampled in increments of five from 10 to 30 to manage computational demands. This analysis (Fig. A.1 in Appendix A) indicated that the range  $k = 15$ –20 provided the most favourable balance between the two metrics. We then conducted a qualitative comparison of models within this range and identified  $k = 18$  as offering the optimal combination of topic coherence and distinctiveness. As the topic model produces a list of topics 1–18 associated with a batch of words, we manually labelled the topics, following the approach of Müller-Hansen et al. [39]. Finding fitting labels involved an assessment of the 50 highest-scoring words for each topic, as well as a qualitative reading of the highest-scoring article per year for each topic. The latter is crucial to understanding the temporal change within topics.

#### 3.2.2. Sentiment analysis

Next, we employ sentiment analysis. A sentiment is defined as an attitude towards a specific topic. We link topic modelling and sentiment in the spirit of Dehler-Holland et al. [49,50] to assess the sentiment associated with topics originating from the topic model. Sentiment analysis matters because political struggle plays out not only over what is spoken about, but over how issues are framed. Tone is central to discursive performance, as actors use it to stage credibility, reinforce the status quo, and marginalise alternatives in the energy transition. In terms of implementation, we improve upon [49] by applying a pre-trained transformer classifier that outputs three probabilities per input token: negative, neutral, and positive. The explicit neutral class is crucial for news text, where much of the fact-based reporting is descriptive. It thus avoids forcing polarity onto language, as is the case in [51]. This pipeline of text mining uncovers sentiments associated with topics along the positivity–neutral–negativity vector



**Fig. 3.** Venn schematic of quantitative methods pipeline. Step 1 (top circle) represents Dynamic Topic Modelling (DTM), which supplies topics. Dashed arrows indicate the sequential flow from DTM to steps 2 and 3, where Sentiment Analysis (2, left circle) and Named Entity Recognition (3, right circle) are applied to the previously identified topics. The intersections indicate the combined quantitative measures (tone-by-topic and actor-by-topic). The text inside each circle summarises the corresponding analytic outputs, see Sections 4–6. This pipeline is complemented by a systematic reading of high-scoring articles to improve depth of analysis.

to analyse news media’s framing of coal in India. Concretely, we use the CardiffNLP Twitter RoBERTa model by Antypas et al. [52].<sup>4</sup> Transformer-based architectures such as RoBERTa have been shown to substantially outperform lexicon-based methods in sentiment classification across diverse domains [53,54], especially for polysemous or domain-specific vocabulary. Furthermore, it provides calibrated class probabilities rather than dictionary-based scores. This makes it particularly suitable for our setting, as for each topic we take the set of FREX-selected stems and score each stem with the classifier.<sup>5</sup> The model’s probabilities are then aggregated per topic using FREX weights to obtain topic-level shares for negative, neutral, and positive sentiment. We also compute a compound, signed score as  $(\text{pos} - \text{neg}) \times (1 - \text{neu})$ , which attenuates high-neutral terms. Although this model is trained on tweets rather than news articles, our application differs from document-level sentiment analysis: we score only FREX-selected stems rather than full news texts. This greatly reduces domain-shift concerns because the classifier is applied to isolated lexical items whose sentiment polarity is relatively stable across genres (see [55] who show that sentiment signals at the lexical/embedding level remain stable across domains). As shown by Rietzler et al. [56], domain-shift issues

<sup>4</sup> The package is publicly accessible via [huggingface.co/cardiffnlp/twitter-roberta-large-topic-sentiment-latest](https://huggingface.co/cardiffnlp/twitter-roberta-large-topic-sentiment-latest) and has been fine-tuned for sentiment analysis on 154 million tweets and evaluated via SuperTweetEval.

<sup>5</sup> To reduce polysemy when scoring unigrams we nudge tokens to a noun reading (e.g., prefixing an article; for known ambiguous items such as “mine” we use fixed noun variants). Further, because legal-investigative vocabulary in our corpus tends to be reported in negative contexts, we incorporate a light domain prior by imposing moderate floors on the negative class for a small set of stems (\*court, case, CBI, order, investig\*, alleg\*, scam\*), with probabilities renormalised; results are robust to reasonable choices of these floors.

in sentiment analysis largely arise from mismatches in contextual and compositional patterns across domains. Because our procedure scores isolated topic-specific stems rather than full news texts, the contextual effects that typically cause domain-transfer errors are largely avoided. As a result, simple lexical sentiment cues transfer more reliably.

### 3.2.3. Named entity recognition

Finally, we use Named Entity Recognition (NER) to identify the actors that populate each topic (*Who* speaks?) (see [57], for a review). The pipeline uses NER to extract people, organisations, governments, and places from articles, structuring otherwise unstructured text. Substantively, NER matters because discursive power is exercised through visibility and voice: who appears in the story space, in which topics, and with what persistence. Mapping actor prominence within topics thus is an attempt to reveal hegemony in practice. We use SpaCy’s pre-trained language model `en_core_web_sm` (2024 release)<sup>6</sup> to tokenize texts and extract named entities from each document, and then aggregate counts by topic for frequency analysis [58]. Before aggregation, we canonicalise aliases and filter non-informative items (for example, the generic “India”) to focus on meaningful actors. SpaCy provides consistent labels (e.g., PERSON, ORG, GPE) and fast, reliable tokenization. Although we use SpaCy’s generic English NER model, we are not classifying documents but extracting only high-salience named entities, and entity surface forms such as major Indian actors (‘Adani’, ‘Coal India’, etc.) are generally well captured even by out-of-domain models. We however manually canonicalise key Indian actors and filter

<sup>6</sup> The package is publicly accessible via [spacy.io/models](https://spacy.io/models), a free open-source library for Natural Language Processing in Python.

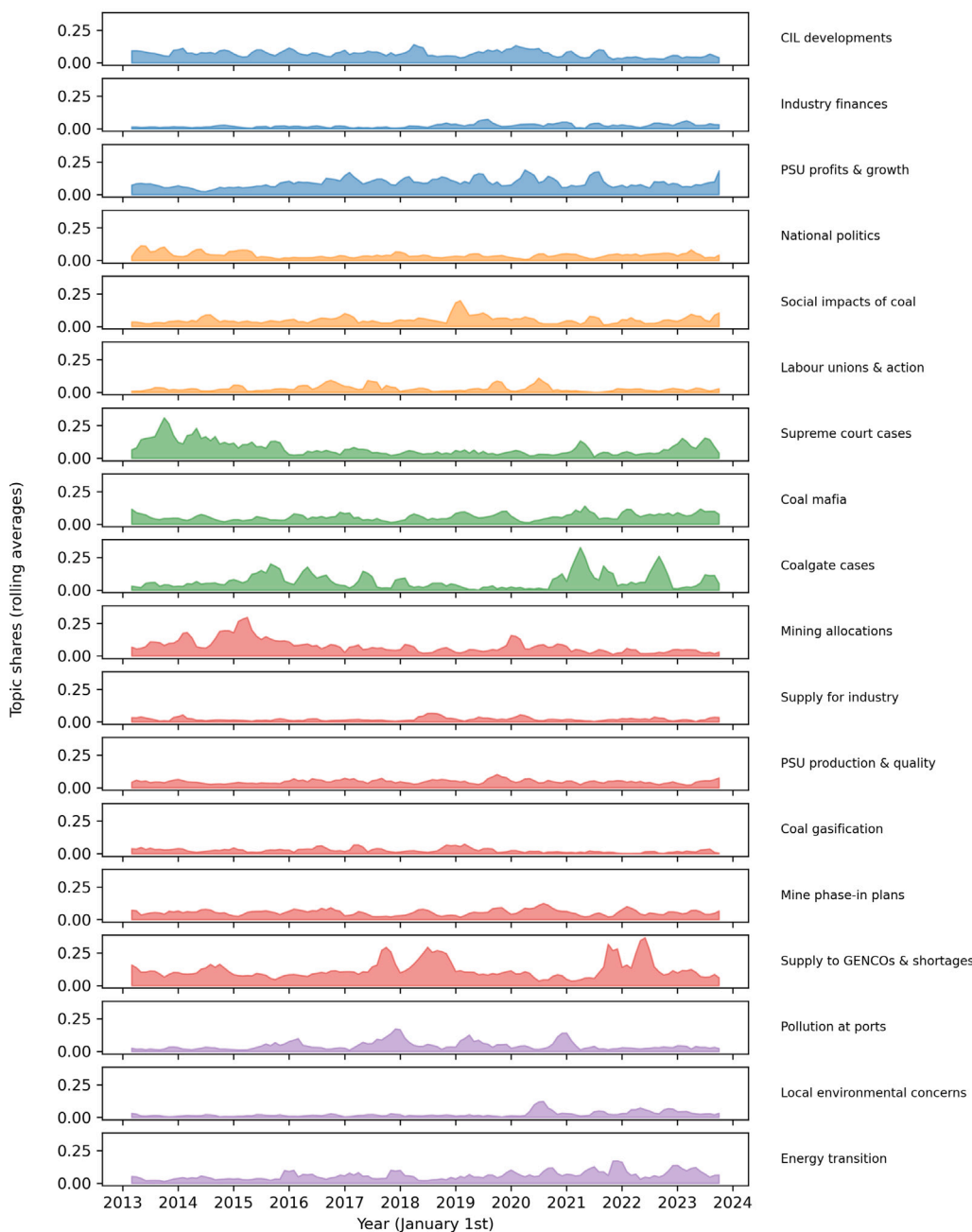


Fig. 4. Graph on four-month rolling averages of 18 individual topic shares over the considered time frame; colour indicates clustering into five narratives.

non-informative entities to overcome the cultural or domain limitations inherent to NER [57,59].

3.2.4. Qualitative analysis of articles

To add a qualitative dimension to the interpretation of the quantitative results, especially the DTM, we conduct a systematic reading of high-scoring articles. For each topic and year, we examine the highest-scoring article, extracting the core argument and narrative, representative quotes, and the stakeholder perspective and category. Where the DTM reveals salient patterns, we increase granularity by extending this sampling to the highest-scoring articles in adjacent time slices (monthly), enabling closer analysis of how arguments and actor perspectives cluster around noteworthy findings. We also embed exemplary quotes from newspapers across Section 4. This step enriches the results sections by grounding the quantitative results in the arguments, voices, and perspectives that shape the actual coal discourse over time.

4. Results on topics, narratives and time trends

Fig. 4 shows four-month rolling averages of topic shares as a result of the DTM. It depicts the share of news articles attributed to each topic in each time slice, which describes the share of the discourse captured by each topic. The considered time frame is on the x-axis. The topic shares as a decimal number representing the percentage of the discourse captured by each topic are on the y-axis, and also show the topic labels.

To synthesise the findings and highlight overarching discursive patterns, the 18 individual topics are grouped into five larger narratives, which are: business, governance, criminality, energy (in-)security and community impacts.

Fig. 5 depicts the temporal evolution of the relative prominence of five narratives. The horizontal axis reports time; the vertical axis reports each narrative's proportional share. Areas are stacked, so the

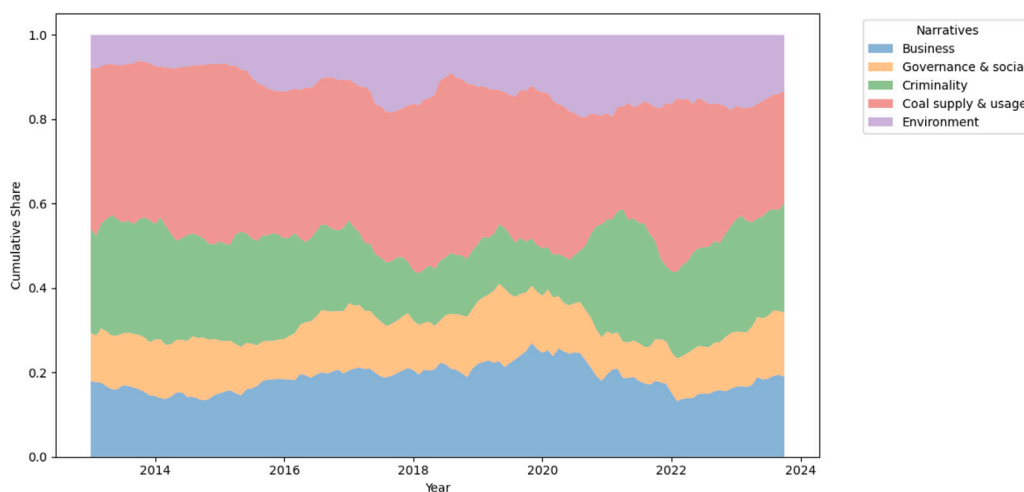


Fig. 5. Graph on five cumulative narrative shares over the considered time frame; colour scheme aligned with Fig. 4. Cumulative share adds up to one by definition.

vertical thickness of a given stratum at time  $t$  represents that narrative's share, and the strata collectively sum to one at each point in time. As shown, the discourse is overwhelmingly structured around coal supply & usage (in red). This includes topics such as supply to GENCOs & shortages, mining allocations, and PSU production & quality (see Fig. 4, also in red), which together present coal as an infrastructural necessity and a cornerstone of national energy security. The business narrative, which includes PSU profits & growth, CIL developments, and industry finances, maintains a stable share throughout the observed period but remains secondary to concerns about energy supply. Criminality (in green), and governance & social (in orange) emerge more episodically, with visible spikes corresponding to investigations, legal interventions, or administrative reforms. However, these narratives rarely challenge the underlying assumption of coal's centrality. Finally, the environment narrative (in purple), covering the topics of pollution at ports, environmental concerns, and energy transition, remains marginal throughout the decade, with only a slight increase in relative share after 2020. This distribution foreshadows the central finding of the analysis: coal is consistently framed as a non-negotiable part of India's development strategy. Even where criticism or calls for reform appear, they are absorbed into a broader narrative of continuity and necessity.

In what follows, we examine the development of these five narratives in more detail (one subsection per narrative), beginning with the prominence of coal supply & usage (in 4.1) and its entanglement with energy transitions, followed by the role of business (in 4.2) as a supporting pillar of continuity, the episodic attention to criminality (in 4.3) and governance & social (in 4.4), and finally the marginal representation of environmental concerns (in 4.5). Each subsection integrates dynamic topic modelling with targeted qualitative reading of representative articles across the topics that constitute each narrative, tracing how salience shifts over time, what receives sustained attention in India's coal discourse, and what remains marginal or episodic.

#### 4.1. Coal supply & usage

The coal supply & usage narrative<sup>7</sup> is the most dominant across the dataset, a reflection of the status-quo. The narrative portrays coal as a strategic necessity and a cornerstone of national development.

<sup>7</sup> This narrative spans six topics: mining allocations, supply to GENCOs & shortages, mine phase-in plans, coal gasification, PSU production & quality, and supply for industry. Generally, consult Fig. 4 to see which topics are contained in each narrative. GENCOs are electric power generation companies, PSU stands for Public Sector Undertaking (a government-owned corporation in India), and CIL stands for Coal India Limited.

Supply to GENCOs & shortages and supply for industry are the most consistently discussed topics throughout the decade. Together with PSU production & quality, these two topics represent the discursive focus on infrastructure efficiency, administrative coordination and operational performance. Such coverage tends to reinforce the perception of coal not as a contested energy source but as a basic requirement whose absence spells crisis.

Further, mining allocations peaks notably (about 30% topic share) in 2014–2015 following the Supreme Court's cancellation of 214 coal block allocations made in the 90's, a scandal commonly known as "Coalgate" (see also 4.3). Beyond this peak, the topic remains prominent, as it reflects the structural importance of ensuring long-term supply security.<sup>8</sup> Mine phase-in plans document the ongoing expansion of coal extraction capacity, both domestically and abroad, in a technical way. Coal gasification similarly appears in a technical register, focusing on technology demonstrations and investment pledges. Gasification is presented as a forward-looking industrial policy, tied to the concept of energy independence and technological progress. In this sense, it mirrors the Indian government's effort to modernise coal rather than transition away from it, exemplified by a new coal gasification financial incentive scheme launched in 2024 with a total outlay of roughly USD 1bn.<sup>9</sup> Technological interventions such as gasification are interpreted not as signs of a possible shift away from coal but as evidence of its evolving centrality. Shortages are treated as logistical crises, not symptoms of deeper structural contradictions.

#### 4.2. Business

The business narrative captures the technical and financial performance<sup>10</sup> of India's coal PSUs (most notably Coal India Limited). Across the full period it remains one of the most stable narratives in the discourse, generally occupying a steady band of coverage with limited long-term variation.

As shown in Fig. 5, the Business narrative exhibits one notable but temporary rise in late 2019 and early 2020, where its share peaks at approximately 24%. This increase stands out relative to its usual range but does not represent a structural break; coverage reduces to its

<sup>8</sup> As noted by Roy and Schaffartzik [14], even after the Court's intervention, the Indian coal complex reasserted itself, and even increased the intensity of extraction.

<sup>9</sup> The scheme aims to achieve 100 million tonnes of coal gasification by 2030, (accessed 03/12/2024).

<sup>10</sup> Articles are largely business reports, tracking production targets, profit margins, and stock prices.

typical level shortly thereafter. The peak appears to reflect a period of intensified routine corporate reporting rather than a fundamental shift in media attention.

Overall, this narrative conveys continuity. Its recurrent presence underscores the administrative and economic centrality of India's coal PSUs. Within [60]'s framework, this coverage reflects what they term the "statecraft coal economy". This refers to a mode of reporting that emphasises formal, institutionalised state control and presents coal as an essential pillar of national development.

#### 4.3. Criminality

The criminality narrative groups together supreme court cases, the Coalgate case, and coal mafia, and captures the legal, political, and criminal entanglements surrounding coal. While not the most dominant narrative by share, it displays clearly defined peaks, particularly in 2014–2015 and again from late 2020 onwards. The first peak (about 29% narrative share in 2014) reflects the aftermath of the Coalgate scandal, culminating in the 2014 Supreme Court ruling that cancelled 214 coal block allocations. The enduring legal fallout is traced in the coverage of special court proceedings and delayed charges, as some issues remained unresolved nearly a decade later.<sup>11</sup> A second wave of prominence (culminating in a peak narrative share of about 31% in 2021) emerges with the 2020 CBI campaign against illegal coal operations, especially in West Bengal. Reporting from this period highlights raids, arrests, and connections to high-ranking political figures. The term "coal mafia" becomes shorthand for a persistent shadow economy surrounding India's coal industry, an economy marked by theft, extortion, and violence. Together, these peaks reveal the deep entanglement of criminality within both the formal and informal coal economies. It is striking that the criminality narrative remains episodic, dominating the discourse only when scandals or crackdowns arise.

#### 4.4. Governance & social

The governance & social narrative consists of national politics, social impacts of coal, and labour unions & action. Together, they account for a fairly modest but consistent share of the discourse. This narrative includes discussion of governance and politics of coal, as well as its consequences on worker and affected communities. Discussions around national politics remains an undercurrent without clear peaks across the timeframe, suggesting that coal is treated less as a political battleground than as a matter of governance. The labour unions & action topic reflects moments of mobilisation, such as the large-scale strikes of 2015 and 2019, as well as around the COVID-19 pandemic in 2020 (peaking at a topic share of about 17%). These strikes often make the national news due to their size and impact. Despite their scale, such labour actions are usually reported in episodic terms and are not accompanied by sustained discussion of long-term labour conditions or structural reform. Though reports about strikes are framed differently depending on the newspaper's political leaning and the origin of the labour action, a common theme is that they are always connected to their impact on energy security. Further, social impacts of coal discuss displacement, environmental health issues, and conflicts over land acquisition. These stories, though often poignant, remain marginal in the overall discourse. They appear as isolated incidents, rarely connected to a systemic critique. For instance: "Earth opens up to swallow man, teenage son in Jharkhand coal mining belt" (Hindustan Times, May 2017). As a whole, this narrative serves more to document events than to shape the dominant coal discourse. Labour insecurity, community

<sup>11</sup> See for instance the quote: "Nearly a decade after cancelling 204 of the 214 coal block allocations since 1993, the SC asked the Central Bureau of Investigation (CBI) [...] why charges had not yet been framed in a majority of cases despite the CBI completing its probe" (The Times of India, 2023).

disruption, and political resistance are present and recurring, but they remain peripheral to the prevailing logic of energy production and economic growth. It is also notable whose stories are almost never told: Participants in the subsistence coal economy, often called "coal pickers", are only mentioned in three out of the 130 top articles for each month.

#### 4.5. Environment

The environment narrative comprises three topics, pollution at ports, local environmental concerns, and energy transition. The first two highlight how environmental harms from coal extraction, transport, and burning are acknowledged in the discourse. For much of the past decade, all topics within this narrative remain sidelined in Indian coal discourse, with only a modest post-2020 increase. Pollution at ports appears most clearly in coverage of coal dust emissions (an acute form of local air pollution) in coastal towns such as Thoothukudi and Goa, responsible for the peaks in 2018 and 2021 (topic shares of about 21% and 19%, respectively). Reports are typically prompted by citizen complaints and protests, legal actions, or state interventions. These stories do clearly identify environmental harms, yet they are reactive, episodic, remain regionalised and thematically isolated. The related local environmental concerns topic includes resistance to coal mining and transport in ecologically sensitive or inhabited areas. Prominent examples include opposition to projects in the Hasdeo Arand forest or along protected coastlines. Such articles discuss issues raised by local communities, activists, or tribal leaders. It is questionable how much space is given to these voices, especially compared to government officials. For instance, the subsequent quote by PM Modi regards the controversy around a new coal mine auction in 2020, which was set to open up vast areas of forest in central India for mining: "Nature has blessed eastern and central India with rich coal deposits. I want to assure people of these regions that labour welfare and development of coal bearing regions are foremost on our mind" (Hindustan Times, June 2020). Together, the topics show that environmental costs are not absent from India's coal discourse but are treated as secondary. They emerge when damage becomes locally visible or politically salient.

We now want to give attention to the topic of energy transition, to come back to our initial question, *From coal to what?*, that motivates our research setup. First, we notice that the energy transition topic's discursive register sets it apart from the two other topics discussed above. It engages with the long-term future of India's energy system, presenting debates around decarbonisation, renewable technologies, and climate policy. Over the ten-year period, this topic gains visible prominence, culminating in a peak in late 2022 at about 16% topic share and sustained coverage into 2023 (compare Fig. 4, last topic, in purple). From qualitative analysis, we see that stronger linkages between coal and climate change appear only from 2015 onwards, coinciding with the growing perception of renewables as a viable complement to fossil fuels. Discussions of a coal phase-out remain limited until 2022, and are driven predominantly by international rather than domestic reports.

While the topic includes a range of framings, from speculative reports on innovative technologies to agenda-setting political commentary, it is primarily shaped by narratives of national interest and energy sovereignty. The 2022 peak corresponds to India's involvement in Just Energy Transition Partnership (JETP) negotiations, following the G20 summit and bilateral announcements involving Indonesia [61]. The articles during this period reflect resistance rather than anticipation, often portraying transition as a matter of geopolitical positioning and financial leverage rather than an environmental necessity. In many cases, the transition is discussed as a global trend that India must navigate on its own terms: gradually, strategically and without compromising development goals.

Throughout the dataset, articles frame the energy transition as a necessary yet distant goal. Immediate policy demands are shaped far more by concerns over energy security than by carbon emissions. As

India's Minister of Power, Raj Kumar Singh, argued in 2019: "The debate about coal needs to be restated or restructured. One should talk about emissions not coal production. This whole premise needs to be re-examined" (Mint, October 2019). There is a stark contrast in recent years between articles arguing for the need of a planned and just transition, and those aligned with the coal complex, which seeks to continue with business as usual for as long as possible. One of the voices that seriously centres the question of a *just* transition, authored by Chandra Bhushan, warns of the risks of repeating past neglect: "So what will happen to these places when coal mining ceases in the next 30–40 years? The truth is we already know what happens when mining stops in a region. [...] All this happened because no long-term plans were put in place to develop an alternative economy" (The Financial Express, November 2018). Warnings from academics and NGOs about the dangers of a future unprepared transition for coal-dependent regions are hence present in the dataset, however, they remain marginal to the mainstream discourse. We suggest that the absence of credible just transition planning is both a policy gap and political choice. Committing to such a transition would signal a fundamental shift in priorities, challenging vested interests and the long-standing alignment between state institutions and coal sector growth. In this light, the energy transition topic reflects a cautious, and often reluctant, engagement with change. It is anchored in national priorities, shaped by developmental concerns, and persistently shadowed by the imperative of energy security. Such marginalisation is reinforced by the media's role as an agenda-setting arena, where elite narratives shape what is considered feasible and urgent.

## 5. Tones across topics

Our contribution lies in demonstrating the analytical value of applying state-of-the-art sentiment analysis to thematically grouped topics derived through topic modelling (see the methods pipeline in Fig. 3). Analysing tone reveals how issues are evaluated, not just discussed, which matters for public acceptance and policy windows in contested transitions. Topic-level tones let us separate short-lived reactions to shocks from durable reframing of the role of coal. Moreover, tone is a key element of discursive performance, through which actors strategically stage credibility and normalise the status quo while marginalising energy alternatives.

By aggregating weighted sentiment probabilities across topics and reporting the composition of negative, neutral and positive, we map the tones embedded in Indian coal discourse, as shown in Fig. 6. We generate the sentiment probabilities with a pre-trained transformer from the Hugging Face Model Hub-CardiffNLP's, as described in the methods subsection (within Section 5). The topics are sorted in descending order by aggregate signed sentiment, with the most positive at the top. The height of the bars on the *x*-axis is normalised to show probability shares that sum to 100%.

Neutral sentiment dominates across all topics (grey bars); the neutral share lies in a tight band of roughly 64%–74%. Positive and negative means are in the 12%–18% and 14%–23% ranges, respectively. At the top end, we find topics tied to system performance and expansion: supply to GENCOs & shortages, energy transition, PSU profits & growth, CIL developments, which exhibit the largest positive shares. On the low end, we find legal–governance topics, which carry the highest negative shares, such as Supreme court cases (negative 23.3%), Coalgate cases (negative 21.4%), followed by local environmental concerns and coal mafia (approx. 17% negative). Most other themes remain ambivalent with differences between positive and negative typically under 5 percentage points, such as energy transition with 16.6% positive and 14.6% negative.

Next, we examine the polarisation of the topics (see Fig. 7), to better show where coverage departs from neutrality and in which direction. The figure plots the weighted positive probability (*x*-axis) against the weighted negative probability (*y*-axis), for each topic. The dashed 45°

line marks parity, such that points below it are relatively more positive, points above it relatively more negative. Topics cluster around the diagonal in the 0.12–0.18 range, indicating a largely neutral baseline in the news coverage. The most polarised topics sit towards the upper right of the plot. Polarised refers to the differences between positive and negative sentiment. As we see, this quadrant is rather empty, suggesting low general polarisation around coal. Supply to GENCOs & shortages (dark blue circle) and energy transition (light blue square) show the clearest dual elevation. PSU profits & growth (orange prism) also sits in this zone.

Interestingly, energy transition appears as the second most positive topic, driven by evaluative and performance-oriented words such as growth, benefit, and progress. We identified these drivers by ranking the words that contribute most to the topic's overall positive tilt. The prominence of words such as project, power, energy, renewables, and clean indicates a framing anchored in business opportunities, technological upgrading, and system expansion. In combination, these drivers situate the transition within an investment and innovation narrative. At the same time, it is among the most polarised topics, with words such as crisis, ambition, decline among the top negative drivers, hence in line with the findings by Shukla and Swarnakar [25].

Generally, the distribution of sentiment points to a bifurcated yet predominantly neutral emotional landscape. Business and operational matters, e.g., supply reliability, and institutional performance, are framed slightly positively, whereas corruption cases, court proceedings, and localised environmental and social harms tend to attract negative tone. Positive framings align with narratives of growth, reliability, and modernisation, bolstering the legitimacy of continued expansion in the coal economy. By contrast, harms are often compartmentalised as legal or local anomalies. They are treated as matters for enforcement or remediation rather than as systemic consequences, often distancing responsibility from decision-makers and from the broader development project (see also Section 4). Our results mirror the nature of news media, where elite discourse simultaneously normalises economic expansion and displaces its adverse consequences onto vulnerable populations. This can steer public debate towards technocratic management of problems rather than substantive consideration of alternatives.

## 6. Actors across topics

The prominence of certain entities shows that it is important who speaks or who is spoken about in political discourse, since narrative dominance structures both visibility and influence. Table 2 below reports the single most frequently mentioned actor per topic, together with its frequency (total count of single entity within the same topic). The order of topics in column 1 is aligned with Figs. 6 and 7. Our goal is to highlight the dominance of certain organisations, agencies, and regions across topics. Indeed, mapping actor prominence across topics reveals how hegemony is exercised in practice. We find that the most frequent entities largely coincide with the industries and institutions central to each thematic area. Coal India Limited consistently appears as the leading actor in the business (equivalent to the business narrative, in blue in Fig. 5) and supply-related discussions (in red in Fig. 5), but interestingly enough also dominates the topic on labour unions and action (with 555 mentions). Unsurprisingly, the Central Bureau of Investigation dominates the legal and corruption-related topics, Supreme court and Coalgate cases. Political debates prominently feature the Congress and the Bharatiya Janata Party, whereas regional indications (Jharkhand, Meghalaya, Goa, and Chhattisgarh) dominate within production, environmental, and social concerns. The energy transition topic stands out in contrast, with China as the most frequent actor, underscoring the international dimension of this debate. China's role is often invoked to justify India's own continued reliance on coal by pointing to relative responsibility, differentiated obligations, and the competitive stakes of decarbonisation (e.g. in terms of industrial competitiveness and geopolitical parity).

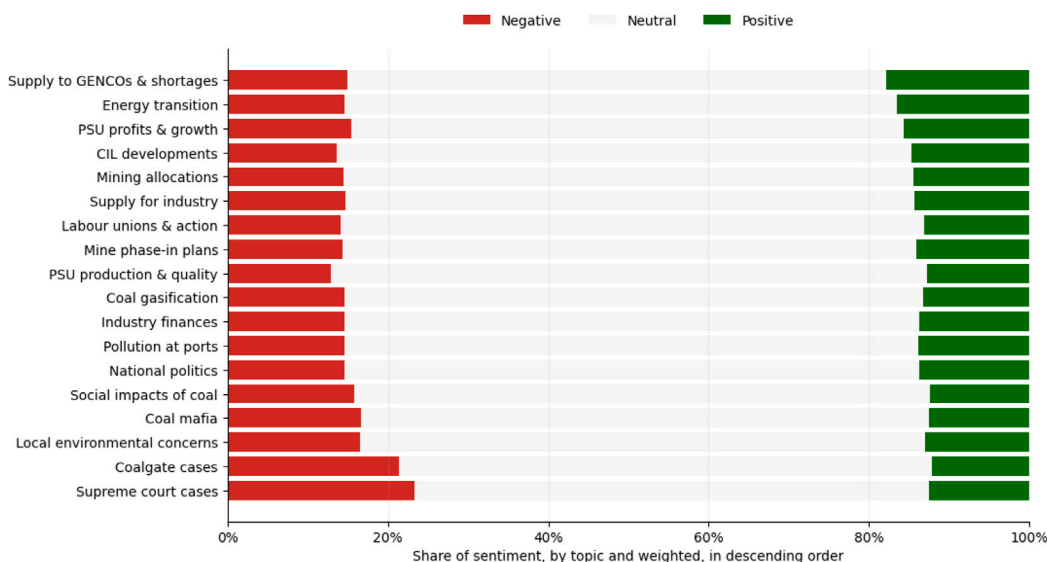


Fig. 6. Sentiment composition by topic, displaying weighted mean probabilities of sentiment scores for 300 FREX words per topic. Negative sentiments are depicted in red, neutral in grey, and positive in green. Topics are sorted in descending order by aggregate signed sentiment, with the most positive at top. Height of bars is normalised to show probability shares that sum to 100%.

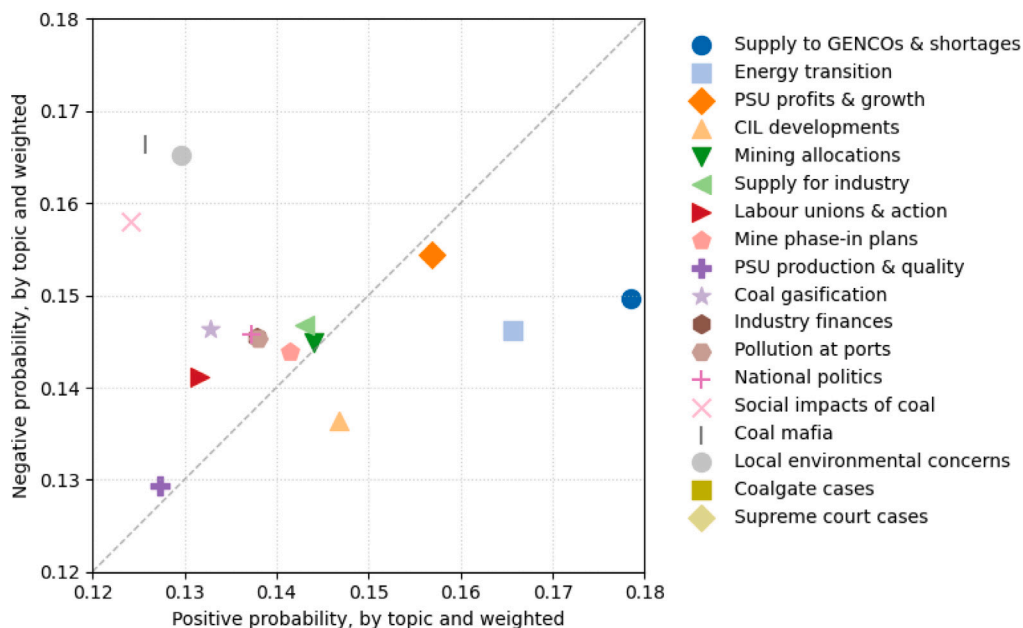


Fig. 7. Topic polarisation scatter plot. Each symbol is a topic. Axes show weighted positive and negative probabilities. The 45° line indicates equal positive and negative tones. Order of topics is in line with Fig. 6.

For more detailed information, Fig. B.1 in Appendix B showcases the topic-specific distributions of the top 15 actors. It highlights that for most topics, one specific actor outnumbers all other actors in terms of frequency. This outlier behaviour aligns with our previous assumption of narrative hegemony and motivates our decision to focus on the most dominant actor in Table 2. Fig. B.1 further shows that recurring entities centre around the key political parties, Indian National Congress (INC) and Bharatiya Janata Party (BJP), and individual politicians such as Piyush Goyal.<sup>12</sup> Similarly, many entities depict major corporations like the National Thermal Power Corporation (NTPC) and Bharat Coking Coal Limited (BCCL), all state-owned coal mining or power generation

companies. Frequent mentions of the coal-rich states Jharkhand, Chhattisgarh, and Odisha reflect regional discourses. Mentions of China, the US, the EU, Glasgow,<sup>13</sup> Australia, and Russia highlight the intricate network of international actors whose coal markets, energy policies, and environmental standards influence and are influenced by India's coal industry.

### 7. Discussion

This paper systematically explores the development of the Indian coal discourse and the state of the discourse around an energy transition

<sup>12</sup> He held several key positions, including the Minister of Railways, Minister of Commerce and Industry, and Minister of Power.

<sup>13</sup> Mentions of Glasgow allude to the climate summit COP26 in 2021, where India opposed the commitment to phase out coal.

**Table 2**

Table of most frequent actor per topic and respective frequency. Order of topics is in line with Figs. 6 and 7.

Topic	Most frequent actor	Frequency
Supply to GENCOs & shortages	Coal India Limited	1076
Energy transition	China	236
PSU profits & growth	Coal India Limited	1410
CIL developments	Coal India Limited	1312
Mining allocations	Coal India Limited	348
Supply for industry	Coal India Limited	137
Labour unions & action	Coal India Limited	555
Mine phase-in plans	Jharkhand	202
PSU production & quality	Coal India Limited	449
Coal gasification	Coal India Limited	592
Industry finances	Coal India Limited	82
Pollution at ports	Goa	175
National politics	Congress	370
Social impacts of coal	Meghalaya	359
Coal mafia	Bharatiya Janata Party	169
Local environmental concerns	Chhattisgarh	214
Coalgate cases	Central Bureau of Investigation	1328
Supreme court cases	Central Bureau of Investigation	2153

away from coal. Our map of long-run narrative dynamics, spanning from 2013 to 2023, seeks to identify where discursive power currently sits. By linking topics, tone, and actors, we illustrate a stable dominance of business and energy-security narratives, with transition talk remaining marginal and typically nested within coal-supply and economic concerns rather than articulated as an autonomous agenda. This finding is comparable with Bez et al. [42] on the state of coal in South Africa and resonates with the argument of Oskarsson et al. [12] and Roy and Schaffartzik [14]. Both show that India's energy policy has evolved towards an energy transition to, rather than away from, carbon-intensive energy based on coal combustion.

Roy and Schaffartzik [14] argue that the persistence of coal reflects underlying power relations, and that these dynamics reproduce injustice even within the expansion of renewables. This argument aligns closely with our findings but operates at a different analytical level: we extend their perspective by focusing explicitly on discursive power. In this landscape, discursive power is held by those actors closest to state capacity and large public sector undertakings, while environmental and governance frames remain in the background. This is reflective of the continued potency of India's national coal economy and the Public Sector Undertakings that are the cornerstone of India's energy security and dominate India's coal complex [8,14]. Our findings echo [20], showing how state-owned incumbents and other powerful actors shape both the narrative and policy space around energy transitions, thus often constraining the visibility of distributive and environmental concerns (see also [21]). Our results are also coherent with Shukla and Swarnakar [2] on the hegemonic narrative that tells the story of necessary coal expansion.

We also find that India's increased efforts to empower the private sector shine through only minorly in the discourse. This implies steady importance of the state, which is in line with Kumar [11] who argues that neoliberal governance has modified the role of the state, as opposed to reduced. Energy security remains a major concern in India as energy demand continues to grow, unsurprisingly peaking around major shortages. Community impacts and complex indigenous rights issues such as the statecraft coal economy in Meghalaya are presented as highly localised issues, a framing that is aligned with India's energy policy, both in terms of its federalised nature and its continued coal dependence [7], and complicates the formation of a larger national resistance movement. The subsistence coal economy is almost invisible in the discourse despite the significant number of workers making a living in this way. Taken together, these patterns illustrate how social and environmental harm is systematically downscaled or localised in ways that reduce their ability to challenge national transition priorities [62]. Another finding is the significant presence of criminality and, more

specifically, major scandals in the discourse, which unsurprisingly have the most negative framing. The result is an agenda architecture oriented towards continuity. Frames that question coal's centrality struggle to sustain attention long enough to shift policy baselines. This is an interesting finding, based on the fact that the pro-coal coalition is dominated by state actors, and the anti-coal coalition by non-state actors [2]. Taken together, this reveals the inability of non-state actors to establish a sustainable counter-discourse. Given that our analysis draws from the most widely read English-language newspapers in the country, these patterns reflect not just elite discourse but the narratives reaching a broad readership. This underscores that our findings hold true not only for policy circles, but also for the information environment available to millions of citizens.

The environment narrative remains marginal across a decade of coal discourse, showing only a modest uptick after 2020. Significant links between coal and climate change emerge only after 2015, when renewables began to be seen as a viable addition to fossil fuels. We also observe a lack of sustained air-pollution coverage despite extensive evidence of its severe and unequal health impacts from coal-fired power generation in India [63]. This suggests that pollution is not framed as a coal-specific issue, limiting its potential influence on energy-transition debates. This notion is corroborated by the lack of pollution discourse in cornerstone papers on India's transition away from coal, such as Chandra et al. [7]. Discourse on a transition involving a phase-out of coal is sparse until 2022 and driven almost exclusively by international reports rather than Indian ones. Business and operational narratives dominate persistently with more positive tone. The coverage of corruption and localised environmental and social harms is framed negatively, on average, though without challenging the underlying assumption of coal's centrality. The debate in general seems to be modestly polarised. Positive frames dominate in the energy transition topic, driven by talk around business opportunities and technological innovation, albeit being one of the more polarised topics. Further, incumbent coal industry stakeholders remain the most prominent actors across the discourse. Their consistent visibility across business, supply, and even labour-related narratives highlights their deep embeddedness in shaping the public conversation. The coal coalition's framing is generally pro-coal rather than anti-transition, seeking instead to muddy the waters regarding the content of a transition ('yes to renewables but also yes to coal'). The same logic can be found in [14], who state that India's logic of "talk renewables, walk coal" constitutes a paradox (see also [7]). In contrast, we do not interpret this pattern as paradoxical, but rather as deliberate and strategic, as ambiguity around transition goals can be intentionally maintained to protect incumbent interests. Further, Shukla and Swarnakar [25] find that renewables attract the

broadest consensus from a discourse coalition perspective to ensure environmental justice, while the necessity to employ fossil fuels to ensure energy access remains a major point of contention.

After the announcement of the JETP with South Africa, a more pronounced Indian anti-coal coalition seems to form, with more articles explicitly discussing a planned phase-out of coal. 2023 is the high point of this development and the only year, in which the word ‘transition’ is among the top 10 topic words. This discursive configuration can help explain why negotiations towards an India JETP ran aground in November 2024. International partners centred the deal on coal retirement and a coal-to-clean pathway. Indian officials and domestic stakeholders consistently prioritised energy security and sovereign control over the transition’s pacing and content. Public reporting confirms that India resisted a coal-focused JETP and preferred cooperation on renewables, grids, storage, and skills.<sup>14</sup> These are areas that align with the dominant domestic frames identified in our analysis, alongside with other concerns such as domestic ownership of transition decisions and the loan-heavy financing structure. To add to this perspective, Chandra et al. [7] further states that the JETP has been a predominantly top-down phenomenon in India.

Our media-based evidence also clarifies why fleeting surges of attention did not translate into durable commitments. Event-driven spikes around environmental accidents or air pollution produce short windows of salience but do not reweigh the overall discourse. Without sustained reframing, these moments clearly fail to open the kind of policy windows needed to anchor a coal retirement-linked finance deal. Similar patterns are visible elsewhere: in South Africa, just transitions have enabled actors to stage justice debates in policy design, but have fallen short of fostering consensus in implementation (Bez et al. forthcoming). Comparative evidence in [65] suggests that countries such as India, where coal dependence remains high, have adopted relatively few policies yet, pointing to an earlier stage of transition. In this light, it may still be premature to expect a structured debate in India about the design of policy instruments. Instead, the discursive field remains oriented around questions of energy security and state capacity, with political economy and sub-national dependence looming as critical challenges for any future policy architecture.

Methodologically, the sequential integration of dynamic topic modelling, sentiment analysis, and named entity recognition yields interpretable, topic-level measures that can be replicated across domains (see also [42]). Substantively, the results delineate where policy attention is most likely to consolidate and how brief windows of heightened salience arise.

Future work could examine links between discursive shifts, elite coalition formation, and observable policy change (see [20]). We also suggest examining Indian news articles that discuss “energy transition” without mentioning coal. Comparing that stream with the coal-centred coverage would show whether transition is framed through different vocabularies, actors, and priorities. Such a contrast could reveal spaces where transition debates unfold on their own terms and where new policy windows might open. Future research could also explore connections to sustainable finance [66] and SDG-oriented strategies [67], assessing how dominant frames around security and modernisation shape the uptake of climate finance or long-term decarbonisation commitments.

This study is not without limitations. First, the absence of a systematic causal identification strategy constrains the ability to draw robust inferences about the drivers of discursive change. Second, the use of dynamic topic modelling entails methodological challenges. As noted by Müller-Hansen et al. [39], it is sensitive to large topic scores in

discrete time windows, which can elevate singular events – such as political scandals like Coalgate – beyond their substantive importance. While this effect appears less pronounced in our analysis due to the shorter time frame, it nonetheless shapes topic prominence. Third, NER does not distinguish whether actors are those actively articulating positions in the discourse or simply being referred to. This ambiguity can complicate conclusions about linguistic roles [59]. A further limitation is our reliance on English-language newspapers, which may underrepresent regional perspectives and frames that are more prominent in vernacular media. Moreover, English-language outlets tend to target urban, more elite readerships, introducing potential biases in how coal-related issues are covered and prioritised (see also [42]). Finally, the positionality of the authors may shape how quantitative outputs are interpreted and how topics are ultimately labelled and framed.

## 8. Conclusion

Taken together, our results show that India’s coal discourse remains structurally anchored in business, supply, and energy-security frames. The energy transition appears only at the margins of public debate. Event-driven surges in attention to environmental or governance issues briefly interrupt this pattern but do not meaningfully alter the underlying narrative architecture. This asymmetry highlights the distinction between episodic shocks and the deeper discursive power of incumbents: while short-term events can momentarily redirect coverage, long-standing coal actors (most visibly Coal India Limited) and influential external players such as China continue to shape the dominant storyline around continuity, reliability, and expansion. By empirically tracing these dynamics over time, our analysis demonstrates how discursive power operates as a stabilising force in energy transition politics, constraining the range of politically viable alternatives. These findings speak to broader questions in the energy-transition literature about how entrenched narratives structure political possibilities. When transition frames are nested within incumbent logics rather than articulated as competing visions, prospects for ambitious coal retirement remain limited. The persistent framing of the transition as modernisation and business opportunity suggests the emergence of a positive – but highly circumscribed – narrative that may only offer support for incremental reform.

From a policy perspective, international and domestic efforts to accelerate India’s energy transition are likely to be more effective when they engage with these existing discursive priorities rather than assume a shared framing of environmental urgency. This implies that energy transition governance is shaped not only by institutional design and policy instruments, but also by the narrative environments in which such instruments are debated and legitimised. Strategies that foreground economic opportunity, just-transition benefits, and energy-security co-benefits may create more durable openings for change, also in light of the JETP failure in 2024. At the same time, strengthening the visibility of environmental and social justice perspectives remains crucial for balancing currently underrepresented voices and for preventing transition discourse from being confined to technocratic or incumbent-friendly frames.

Future research could extend this work by analysing multilingual media ecosystems, incorporating social-media dynamics, or conducting cross-national comparisons between JETP (aspiring) countries to assess how discursive configurations vary across coal-dependent contexts. Such work would help clarify how narrative dominance interacts with institutional contexts to shape energy transition governance outcomes, and would deepen our understanding of how narrative structures shape transition politics and where meaningful opportunities for reframing may arise.

<sup>14</sup> See for instance [www.context.news/just-transition/india-wants-a-clean-energy-transition-deal-and-its-coal-too](http://www.context.news/just-transition/india-wants-a-clean-energy-transition-deal-and-its-coal-too) and [www.cleanenergywire.org/news/india-donor-countries-give-just-energy-transition-partnership-german-official](http://www.cleanenergywire.org/news/india-donor-countries-give-just-energy-transition-partnership-german-official) (both accessed 15/08/2025).

**CRedit authorship contribution statement**

**Arne Arens:** Conceptualization, Methodology, Software, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Charlotte Sophia Bez:** Conceptualization, Methodology, Software, Formal analysis, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgements**

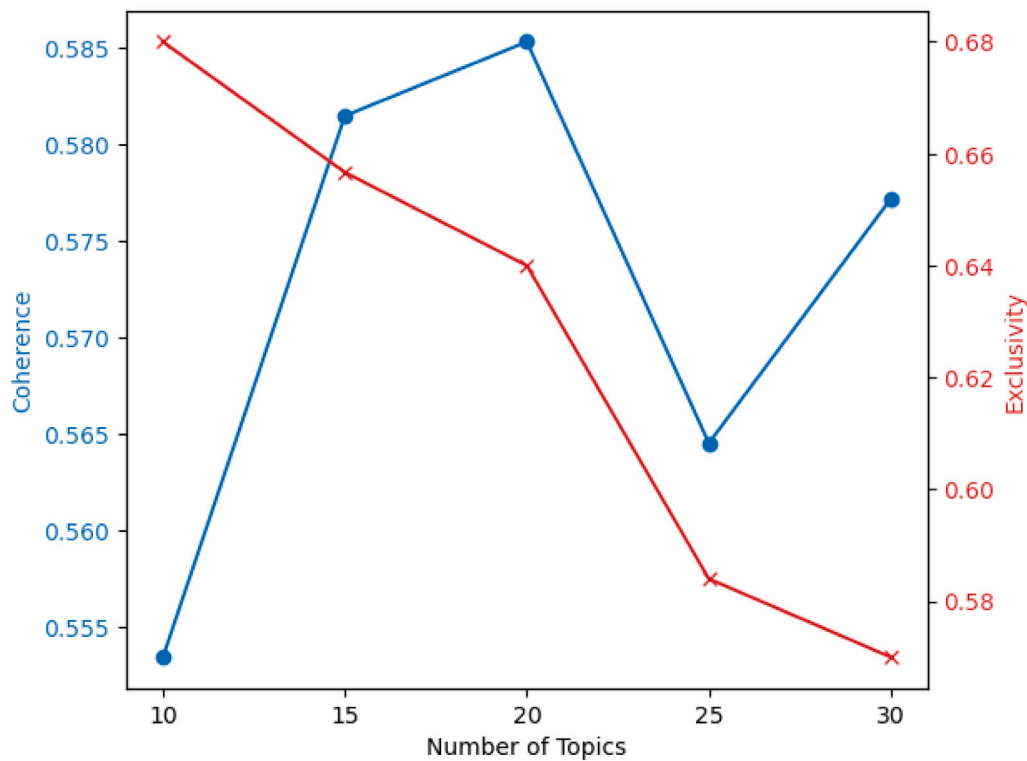
C. S. Bez acknowledges the financial support by the Werner Siemens Foundation and funding through the TRANSFORM project by the German Federal Ministry for Economic Cooperation and Development.

**Appendix A**

See Figs. A.1–A.3 and Table A.1.

**Appendix B**

See Fig. B.1.



**Fig. A.1.** Coherence (in blue) and Exclusivity scores (in red) for the sampled k (10–30).

**Table A.1**

Table of Indian Readership Survey (IRS) statistics for five different years (2013–2019) for all news outlets in the dataset.

Name of news outlet	IRS 2013	IRS 2014	IRS 2017	IRS 2019	IRS Business dailies 2019
The Times of India	1st	1st	1st	1st	–
Hindustan Times	2nd	2nd	2nd	3rd	–
The Hindu	3rd	3rd	3rd	2nd	–
The Telegraph	5th	5th	7th	9th	–
The Economic Times	6th	6th	4th	4th	1st
The Indian Express	–	–	6th	5th	–
The New Indian Express	–	–	8th	6th	–
Mint	–	–	–	–	2nd
Business Standard	–	–	–	–	3rd
The Hindu Business Line	–	–	–	–	4th
The Financial Express	–	–	–	–	5th

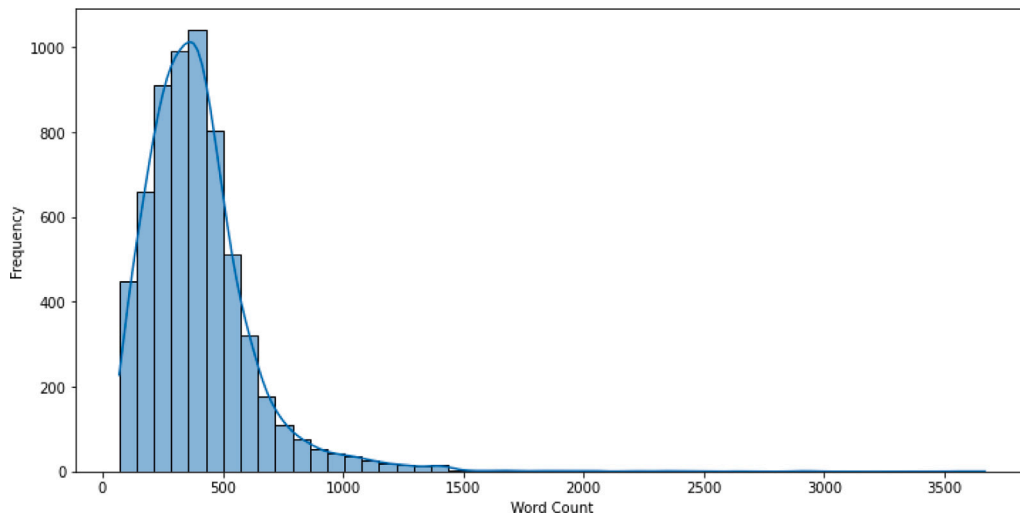


Fig. A.2. Distribution of length (word count) of the 6299 news articles.

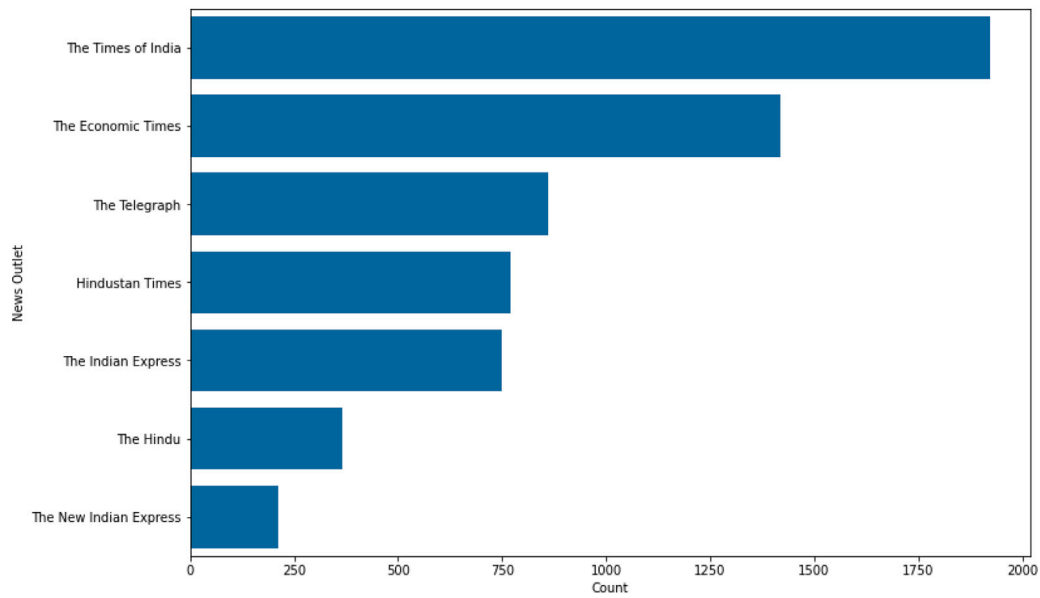


Fig. A.3. News outlet distribution of the 6299 news articles, sorted by count.

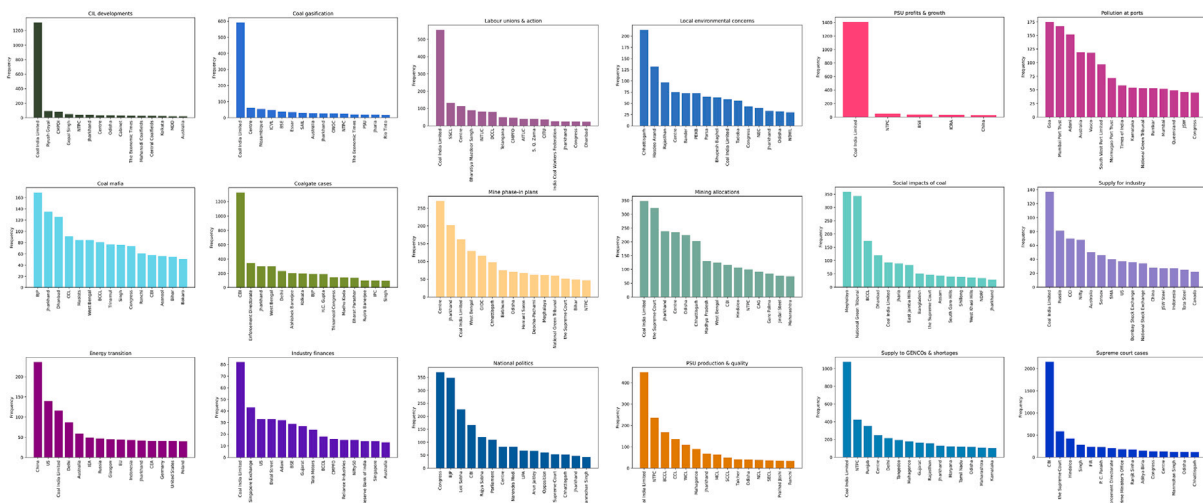


Fig. B.1. Top 15 entities per topic, sorted by frequency and based on Named Entity Recognition (NER).

## Data availability

The code used in this study is publicly available via Zenodo (DOI: <https://doi.org/10.5281/zenodo.18348399>). Due to licensing restrictions, the primary news text data obtained from LexisNexis (lexis\_india\_coal.json) cannot be shared publicly.

## References

- [1] B. Blankenship, M. Aklin, J. Urpelainen, V. Nandan, Jobs for a just transition: Evidence on coal job preferences from India, *Energy Policy* 165 (2022) 112910.
- [2] R. Shukla, P. Swarnakar, Energy transition and dialectics: tracing discursive resistance to coal through discourse coalition in India, *Globalizations* 20 (8) (2023) 1296–1311.
- [3] S.S. Alharbi, M. Al Mamun, S. Boubaker, S.K.A. Rizvi, Green finance and renewable energy: A worldwide evidence, *Energy Econ.* 118 (2023) 106499.
- [4] Press Information Bureau (PIB), Government of India, India's RE capacity registers 15.84% Year-on-Year growth: Record capacity addition in solar and wind; solar at 97.86 GW, wind at 48.16 GW; total RE installed capacity reaches 209.44 GW, 2025, . (Accessed 17 February 2025).
- [5] Central Electricity Authority, All india installed capacity (in MW) of power stations, 2023, URL [cea.nic.in/wp-content/uploads/installed/2023/06/IC\\_June\\_2023\\_Updated.pdf](http://cea.nic.in/wp-content/uploads/installed/2023/06/IC_June_2023_Updated.pdf). (Last Retrieved 17 February 2025).
- [6] V. Sharma, J. Loginova, The social aspects of India's energy transition: Coal, critical minerals and Socio-Economic dependencies, in: *The Role of Coal in a Sustainable Energy Mix for India*, Routledge India, 2023, pp. 75–103.
- [7] R. Chandra, S. Pai, S. Nayak, S.H. Devagudi, India's coal conundrum: Decarbonization amidst a developmental legacy, *Wiley Interdiscip. Rev.: Clim. Chang.* 16 (1) (2025) e928.
- [8] K. Lahiri-Dutt, The diverse worlds of coal in India: Energising the nation, energising livelihoods, *Energy Policy* 99 (2016) 203–213.
- [9] S. Rosewarne, The transnationalisation of the Indian coal economy and the Australian political economy: The fusion of regimes of accumulation? *Energy Policy* 99 (2016) 214–223.
- [10] K. Kohli, M. Menon, The Tactics of Persuasion: Environmental negotiations over a corporate coal project in coastal India, *Energy Policy* 99 (2016) 270–276.
- [11] M. Kumar, Fossil neoliberalism and its limits: Governing coal in South India, *Environ. Plan. E: Nat. Space* 5 (4) (2022) 1853–1871.
- [12] P. Oskarsson, K.B. Nielsen, K. Lahiri-Dutt, B. Roy, India's new coal geography: Coastal transformations, imported fuel and state-business collaboration in the transition to more fossil fuel energy, *Energy Res. Soc. Sci.* 73 (2021) 101903.
- [13] M. Kumar, Violent transitions: towards a political ecology of coal and hydropower in India, *Clim. Dev.* 16 (9) (2024) 751–761.
- [14] B. Roy, A. Schaffartzik, Talk renewables, walk coal: The paradox of India's energy transition, *Ecol. Econom.* 180 (2021) 106871.
- [15] Y.-E. Park, Research evidence for reshaping global energy strategy based on trend-based approach of big data analytics in the corona era, *Energy Strategy Rev.* 41 (2022) 100835.
- [16] P. Vishnuram, S. Alagarsamy, M. Bajaj, M. Alqahtani, I. Ahmed, M. Khalid, A critical review on alternative fuels for road transportation: Pollution mitigation, environment and economic perspective, *Energy Strategy Rev.* 61 (2025) 101894.
- [17] M. Aperi, L. Eicke, A. Goldthau, M. Hashem, S. Huneus, R.L. de Oliveira, M. Otieno, E. Schuch, K. Veit, An energy justice index for the energy transition in the global south, *Renew. Sustain. Energy Rev.* 192 (2024) 114238.
- [18] M.D. Jones, M.K. McBeth, A narrative policy framework: Clear enough to be wrong? *Policy Stud. J.* 38 (2) (2010) 329–353.
- [19] E.A. Shanahan, M.D. Jones, M.K. McBeth, Policy narratives and policy processes, *Policy Stud. J.* 39 (3) (2011) 535–561.
- [20] B.K. Sovacool, M.-C. Brisbois, Elite power in low-carbon transitions: A critical and interdisciplinary review, *Energy Res. Soc. Sci.* 57 (2019) 101242.
- [21] H.-M. Husu, Rethinking incumbency: Utilising Bourdieu's field, capital, and habitus to explain energy transitions, *Energy Res. Soc. Sci.* 93 (2022) 102825.
- [22] A. Pattison, W. Cipolli III, J. Marichal, The devil we know and the angel that did not fly: An examination of devil/angel shift in twitter fracking "debates" in NY 2008–2018, *Rev. Policy Res.* 39 (1) (2022) 51–72.
- [23] M.D. Jones, G. Song, Making sense of climate change: How story frames shape cognition, *Political Psychol.* 35 (4) (2014) 447–476.
- [24] T. Bolsen, M.A. Shapiro, The US news media, polarization on climate change, and pathways to effective communication, *Environ. Commun.* 12 (2) (2018) 149–163.
- [25] R. Shukla, P. Swarnakar, Energy justice in post-Paris India: Unpacking consensus and conflict through storylines and discourse coalitions, *Energy Res. Soc. Sci.* 91 (2022) 102687.
- [26] W. Conway-Lamb, Pluralizing climate change adaptation: mapping discourses in Vietnam, *Clim. Dev.* (2024) 1–12.
- [27] N. Joshi, N. Ranjit, A. Peddibhotla, C. Singh, Lost and damaged: a systematic review of current loss and damage due to climate change in India, *Clim. Dev.* (2025) 1–15.
- [28] S.S. Vishwanathan, P. Fragkos, K. Fragkiadakis, A. Garg, Assessing enhanced NDC and climate compatible development pathways for India, *Energy Strategy Rev.* 49 (2023) 101152.
- [29] K.A. Khan, A. Anjum, M. Subhan, N.A.B. Khairuddin, Evaluating the disaggregated impact of renewable and non renewable electricity generation on economic welfare in India, *Sci. Rep.* 15 (1) (2025) 26725.
- [30] A. Mallett, P. Pal, Green transformation in the iron and steel industry in India: Rethinking patterns of innovation, *Energy Strategy Rev.* 44 (2022) 100968.
- [31] J. Yang, J. Urpelainen, The future of India's coal-fired power generation capacity, *J. Clean. Prod.* 226 (2019) 904–912.
- [32] X. Liu, Z. Liu, Exploring the impact of media discourse on social perceptions towards biomass energy utilization in China, *Energy Strategy Rev.* 42 (2022) 100896.
- [33] A. Jungherr, O. Posegga, J. An, Discursive power in contemporary media systems: A comparative framework, *Int. J. Press/Politics* 24 (4) (2019) 404–425.
- [34] J.A. Ordonez, M. Jakob, J.C. Steckel, H. Ward, India's just energy transition: Political economy challenges across states and regions, *Energy Policy* 179 (2023) 113621.
- [35] Y. Liu, Bridging research and policy in China's energy sector: A semantic and reinforcement learning framework, *Energy Strategy Rev.* 59 (2025) 101770.
- [36] M. Zhang, T. Kober, A. Patt, E. Panos, Mapping policy priorities across diverse stakeholder groups in the swiss energy and climate subsystems, *Energy Strategy Rev.* 61 (2025) 101830.
- [37] J. Kuenzler, B. Stauffer, C. Schlauffer, G. Song, A. Smith-Walter, M.D. Jones, A systematic review of the Narrative Policy Framework: a future research agenda, *Policy Politics* (2024) 1–23.
- [38] L.L. Benites-Lazaro, L. Giatti, A. Giarolla, Topic modeling method for analyzing social actor discourses on climate change, energy and food security, *Energy Res. Soc. Sci.* 45 (2018) 318–330.
- [39] F. Müller-Hansen, M.W. Callaghan, Y.T. Lee, A. Leipprand, C. Flachsland, J.C. Minx, Who cares about coal? Analyzing 70 years of German parliamentary debates on coal with dynamic topic modeling, *Energy Res. Soc. Sci.* 72 (2021) 101869.
- [40] P. Giganti, C. Barra, P.M. Falcone, Unveiling the nexus of institutional quality and renewable energy utilizing a Topic Modelling approach, *Renew. Sustain. Energy Rev.* 214 (2025) 115516.
- [41] B. Jiang, M.Y. Raza, Research on China's renewable energy policies under the dual carbon goals: A political discourse analysis, *Energy Strategy Rev.* 48 (2023) 101118.
- [42] C.S. Bez, G. Raederscheidt, J.C. Steckel, Dirty talk: Media discourse and the struggle over South Africa's coal transition, *Energy Res. Soc. Sci.* 130 (2025) 104398.
- [43] A.A. Guenduez, T. Mettler, Strategically constructed narratives on artificial intelligence: What stories are told in governmental artificial intelligence policies? *Gov. Inf. Q.* 40 (1) (2023) 101719.
- [44] A. Sen, D. Ghatak, G. Khanuja, K. Rekha, M. Gupta, S. Dhakate, K. Sharma, A. Seth, Analysis of media bias in policy discourse in india, in: *Proceedings of the 5th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies, 2022*, pp. 57–77.
- [45] Media Bias/Fact Check, Media bias and credibility ratings, 2024, URL [mediabiasfactcheck.com](http://mediabiasfactcheck.com). (Accessed 06 December 2024).
- [46] D.M. Blei, J.D. Lafferty, Dynamic topic models, in: *Proceedings of the 23rd International Conference on Machine Learning, 2006*, pp. 113–120.
- [47] P. DiMaggio, M. Nag, D. Blei, Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding, *Poetics* 41 (6) (2013) 570–606.
- [48] A. Törnberg, P. Törnberg, Muslims in social media discourse: Combining topic modeling and critical discourse analysis, *Discourse Context. Media* 13 (2016) 132–142.
- [49] J. Dehler-Holland, M. Okoh, D. Keles, Assessing technology legitimacy with topic models and sentiment analysis—The case of wind power in Germany, *Technol. Forecast. Soc. Change* 175 (2022) 121354.
- [50] J. Dehler-Holland, K. Schumacher, W. Fichtner, Topic modeling uncovers shifts in media framing of the German renewable energy act, *Patterns* 2 (1) (2021).
- [51] J. Hartmann, M. Heitmann, C. Siebert, C. Schamp, More than a feeling: Accuracy and application of sentiment analysis, *Int. J. Res. Mark.* 40 (1) (2023) 75–87.
- [52] D. Antypas, A. Ushio, F. Barbieri, L. Neves, K. Rezaee, L. Espinosa-Anke, J. Pei, J. Camacho-Collados, Supertweeteval: A challenging, unified and heterogeneous benchmark for social media nlp research, 2023, arXiv preprint arXiv:2310.14757.
- [53] L. Zhang, S. Wang, B. Liu, Deep learning for sentiment analysis: A survey, *Wiley Interdiscip. Rev.: Data Min. Knowl. Discov.* 8 (4) (2018) e1253.
- [54] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, Roberta: A robustly optimized bert pretraining approach, 2019, arXiv preprint arXiv:1907.11692.
- [55] T. Mikolov, I. Sutskever, K. Chen, G.S. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, *Adv. Neural Inf. Process. Syst.* 26 (2013).
- [56] A. Rietzler, S. Stabinger, P. Opitz, S. Engl, Adapt or get left behind: Domain adaptation through BERT language model finetuning for aspect-target sentiment classification, in: *Proceedings of the Twelfth Language Resources and Evaluation Conference, 2020*, pp. 4933–4941.

- [57] A. Goyal, V. Gupta, M. Kumar, Recent named entity recognition and classification techniques: a systematic review, *Comput. Sci. Rev.* 29 (2018) 21–43.
- [58] H. Shelar, G. Kaur, N. Heda, P. Agrawal, Named entity recognition approaches and their comparison for custom ner model, *Sci. Technol. Libr.* 39 (3) (2020) 324–337.
- [59] B. Jehangir, S. Radhakrishnan, R. Agarwal, A survey on named entity recognition—datasets, tools, and methodologies, *Nat. Lang. Process. J.* 3 (2023) 100017.
- [60] K. Lahiri-Dutt, D.J. Williams, The coal cycle: small-scale illegal coal supply in eastern India, *J. Resour. Energy Dev.* 2 (2) (2005) 93–105.
- [61] A. Seiler, H. Brown, S. Matthews, The JETPs of South Africa and Indonesia, *Cent. Glob. Dev.* (2023).
- [62] D. Schlosberg, Reconceiving environmental justice: global movements and political theories, *Environ. Politics* 13 (3) (2004) 517–540.
- [63] J. Kopas, E. York, X. Jin, S. Harish, R. Kennedy, S.V. Shen, J. Urpelainen, Environmental justice in India: incidence of air pollution from coal-fired power plants, *Ecol. Econom.* 176 (2020) 106711.
- [64] C.S. Bez, T. Klug, J. Steckel, G. Raederscheidt, Converging visions, contested narratives: media discourse analyses of South Africa’s just transition, in: M. Pansera, C. Vázquez Santos (Eds.), *ESEE-Degrowth 2024: Science, Technology, and Innovation beyond growth. Cultivating collective creativity for a sustainable future*, 2024.
- [65] P. Yanguas Parra, A. Furnaro, N. Malz, L. Wijayani, P.-Y. Oei, Supply-side policies in coal transitions and country vulnerabilities: a systematic literature review, *Environ. Politics* (2025) 1–37.
- [66] R.R. Bhandary, K.S. Gallagher, F. Zhang, Climate finance policy in practice: A review of the evidence, *Clim. Policy* 21 (4) (2021) 529–545.
- [67] A.M. Khalid, S. Sharma, A.K. Dubey, Concerns of developing countries and the sustainable development goals: Case for India, *Int. J. Sustain. Dev. World Ecol.* 28 (4) (2021) 303–315.