

Villainous or Valiant? Depictions of Oil and Coal in American Fiction and Nonfiction Narratives

Abstract Energy is fundamental to human existence, and its ubiquity has allowed energy to enter cultural consciousness in a manner that is reflected in the stories that we tell ourselves, and others. Modern society runs on fossil fuels like oil and coal, two resources that are frequently discussed in part due to their contribution to both positive and negative outcomes. This research uses a digital ecocritical approach to explore a corpus of 60 narratives, both fictional and non-fictional, published between 2002 and 2016 by US authors. We combine text mining methods including sentiment analysis and topic modeling with selected manual review of texts to posit that American narratives often depict oil as new and exciting, with associated dangers seen as tragic but thrilling. Appalachian coal, by contrast, is portrayed nostalgically, depicted as a nearly familial presence that has betrayed its communities and no longer represents security and prosperity. Thus, while oil is hypothetical and exciting, coal is real and disappointing. Latent cultural attitudes about these and other resources can provide insights as to how Americans view ongoing deployment of energy infrastructure. Further, understanding cultural context can help direct attention to issues of high significance to communities experiencing energy development.

Keywords: fossil fuels, narrative, text mining, digital humanities

1 Introduction

Modern industrial energy resources are increasingly visible in our communities, on our landscapes, and in our concerns about environmental quality (see e.g. [1-4]). The effects of using modern fossil fuels are felt as climate change, other forms of pollution, economic growth and contraction, and societal change. The infrastructure of these energy systems is everywhere—from mines, wells, power lines, pipelines, power plants, refineries, vehicles, to the outlets in our homes. The stories that we tell reflect our culture and, whether consciously or not, encode our opinions, thoughts and beliefs [5-7]. Recognizing this chain of influence, we posit that the ubiquity of energy infrastructure in the United States has contributed to the introduction of latent attitudes about energy to written narratives. An analysis of these narratives, then, can reveal information about society's view of resources that continue to be deployed in communities.

This research investigates novels and narrative nonfiction aimed at nonspecialist audiences to explore how contemporary American narratives depict two major fossil energy resources with long histories in the United States: oil and Appalachian coal. We combine digital methods with an ecocritical approach [8], specifically by using resource-oriented texts to “examin[e] the importance of environmental values” [7]. Specifically, we analyze narrative treatment of two fossil fuel energy resources in recent US fiction and nonfiction, focusing on latent assumptions about oil and coal, how depictions of each resource change over the course of a narrative, the role of each resource as character or setting, and the emotional valences of description of oil and coal resources and communities. We add this investigation of how cultural understanding of oil and Appalachian coal is encoded in recent published texts, whether explicitly foregrounded in major plot elements or simply as a texture in the fabric of a narrative, to other work addressing the relationship between energy and cultural attitudes (see e.g. [9] on oil culture and [10] on environmental justice in late 19th/early 20th century novels).

Our corpus consists of 60 texts written by American authors since 2002 (15 each for coal and oil fiction and narrative nonfiction). This geographically and temporally resolved snapshot of conventional energy-related literature enables us to assess common associations with oil and Appalachian coal. Noting the observed historical influence of literature on environmental policy [11], we motivate this work by recognizing that latent assumptions about each resource might affect policy design, implementation, and evaluation in the United States, particularly related to issues like greenhouse gas controls and differential treatment of the oil-dominated transportation sector versus the coal-dominated electricity sector in the face of climate change.

1.1 The Oil Corpus

Understanding the recent history of the American oil industry offers a useful context for the contemporary corpus selected for this investigation. In particular, recognizing the importance of high volume, multi-stage hydraulic fracturing of horizontal wells and deepwater offshore drilling to recent US oil production is relevant for understanding the corpus. These two developments have changed the industry rapidly and recently, most notably by substantially expanding US oil production [12]. Neither offshore drilling nor hydraulic fracturing are new in the US: the first US offshore well was drilled in the late 1800s [13], Ch. 12, while hydraulic fracturing in vertical wells began in the late 1940s [14], Ch. 1. However, deployment of these concepts for what is called unconventional production—in deep water and in high volume applications to horizontal wells in tight formations, respectively—has fundamentally altered American oil production. This article speaks to this unconventional application and the changes it has effected on the physical and cultural landscape.

Both hydraulic fracturing and offshore drilling are well represented in the corpus: crucially, each is associated with both economic security and environmental concern. The rapid increase in American oil production during the period represented by the corpus means that the US' dependence on foreign countries for oil has changed substantially from the beginning of the corpus period to the present—a relevant change in a country where the oil crises of the 1970s [15] are still within living memory for many adults. Also, importantly, both hydraulic fracturing and offshore drilling have been matters of public concern and fear. For hydraulic fracturing, this concern has manifested largely in the form of environmental concerns related to water contamination and other issues [16]. Both safety and environmental concerns about offshore drilling were amplified by the dramatic 2010 Gulf of Mexico BP offshore oil spill on the Deepwater Horizon rig after a blowout at the Macondo prospect killed 11 and became one of history's largest oil spills [17, 18].

While hydraulic fracturing and offshore drilling emerged in response to the same economic conditions, namely the rise in oil prices starting around 2000 [19], the two technologies are personalized to different degrees. That is, the connection between resource extraction and what it means for an individual is quite different. Hydraulic fracturing has led to a proliferation of wells that are near people, frequently visible, and often associated with individual wealth creation. Those who own land and mineral rights relevant for development often reap large financial rewards, resulting in a phenomenon sometimes called “shaleionaires” referring to situations where individuals suddenly become millionaires from exploitation of oil or natural gas shale resources (for examination of this phenomenon in gas fields, see e.g. [20]). Many concerns are therefore local, including increased traffic, changes to community character, and worry about

water and air quality effects [21]. Similarly, communities near offshore drilling activities evidence concerns about issues like oil spills and worker safety [22, 23]. Communities with people who work offshore are affected by the rhythms of offshore work, where workers are often gone, then home, for weeks at a time [24, 25]. However, offshore drilling results in wells that are far bigger, far more expensive, and far less likely to create financial windfalls for individuals than wells enabled by hydraulic fracturing.

1.2 The Coal Corpus

In contrast to the recent expansion of oil production in the US, the recent history of Appalachian coal has been one of decline. Appalachian coal production has fallen by about 40% since 2000 [26]. Demand for high quality coal for steelmaking produced a mini-boom in Appalachia in 2008, but this boom was a short-lived precursor to the rapid 30% drop in production observed since 2008. Layoffs and bankruptcies have followed.

Appalachia has a long history with coal, including disputes about land tenure, environmental damage, and labor activism [27-31] to more recent environmental activism such as 2009's Mountain Justice Summer, centered on mountaintop removal coal mining [32, 33]. This deeply place-rooted and often activist history is present in cultural memory through heritage designations and reenactments [31]. The long and contentious history of collective action in Appalachian coal communities is an important point of difference with the history of American oil production, where unions and strikes have existed but are much less central to oil's narrative. This difference is perhaps related to oil's shorter global history, US mineral ownership practices, and realities of infrastructure scale that allow individuals to strike out on their own and occasionally succeed; this is something that is not really possible in the coal industry.

In the selected corpus, destruction of place relationships is an overarching theme in coal depictions, focused on two major past-oriented pathways: the disappearance of jobs leading to community decline and the highly visible environmental destruction associated with mountaintop removal. Both are consequence-oriented topics associated with lived experience of economic and environmental harm. Focus areas in individual texts include specific (and usually negative) events, like massacres associated with union busting or the opening or closure of a particular mine. Notably, the ostensibly technology related topic of mountaintop removal mining, a relatively new technique for coal extraction in Appalachia, manifests in the corpus as an environmental justice and social upheaval topic, not as a technological topic. Mountaintop removal texts in the corpus tend to focus on consequences rather than opportunity.

2 Methods

2.1 Corpus selection

The corpus of contemporary American oil- and coal-related texts is nonrandomly selected based on triangulation from compiled libraries of Appalachian fiction, bibliographies of industry narratives, award-winning works addressing oil or Appalachian coal, and texts recently encountered by the authors as part of a broader study of attitude formation in energy communities. This core is supplemented by a broad search of Amazon's Kindle e-book library for narratives whose titles and short descriptions include references to terms like "oil" and "coal": this expands the corpus to include less acclaimed works, including romance novels, self-published works, and others. While the corpus is nonrandomly selected, the size of the corpus

was chosen to be sufficiently large that further expansion in all four subcorpora—oil fiction, oil nonfiction, coal fiction, and coal nonfiction—would be challenging given search constraints of publication after 2000 (in practice, the corpus reflects texts published in 2002 and later), US authorship, and narrative format. As text mining techniques are often adept at identifying authorship (e.g. [34-36]), sometimes to the exclusion of other characteristics of the text [37], the corpus was also selected to avoid author replication. Care was taken to ensure that each subcorpus is not dominated by a single genre: for example, thrillers that reference oil are relatively common. Finally, texts were examined to ensure that either coal or oil is important to the overall narrative.

The final corpus includes 60 narrative texts published since 2002 by American authors, divided into four subcorpora of 15 texts each, categorized as American oil-related fiction, American oil-related nonfiction, Appalachian coal-related fiction, and Appalachian coal-related nonfiction (Table S1). Oil-related texts are overall newer than coal-related texts, with the earliest included publication dates in 2006 (nonfiction) and 2008 (fiction) for oil compared with 2004 (nonfiction) and 2002 (fiction) for coal. Four oil texts were published in 2016, versus none for coal. This discrepancy possibly reflects search bias, but effort was made to select broadly. Notably, the selected time period covers major disaster events (BP's Deepwater Horizon oil spill and Massey's Upper Big Branch mining disaster), price changes, and emergence of public concern about extraction methods for both resources.

2.2 Analytical approach

2.2.1 Non-digital approach

To guide analysis and provide a basis for sense checking outputs from digital analyses, twenty-one of the sixty texts in the corpus were read in full. This manual review extends one of the authors' direct experience through fieldwork in Appalachian coal communities and American oil communities as useful context on the content of the corpus. The other author's expertise in literary theory, narrative structure, and literary interpretation through digital humanities also provides a strong basis for assessing and interpreting digital results. Human judgment in defining research questions, selecting the corpus and analyses, and interpreting results is supported by systematic, replicable digital analysis that exceeds the human authors' capacity to quickly and consistently evaluate large amounts of data and identify micro and macro patterns [38].

2.2.2 Digital approach

In both the social sciences and the humanities, computational analysis of textual data has gained increasingly widespread acceptance as a set of methods uniquely suited to large-scale historical and cultural analysis [39, 40]. These techniques allow researchers to both analyze textual data at very large scales, as well as gather fine-grained information on specific word usage, that may fall under the threshold of comprehension for both readers and authors [41]. In working across these scales, the methods that we use here (topic modeling and sentiment analysis) simultaneously offer the ability to compare high-order word patterns across texts, as well as trace subtle textual cues within individual narratives [42]. Both methods are particularly well suited for the corpus that we consider here. Topic models combine probabilistic modeling with subjective, human-based, interpretive choices, allowing us to prime our statistical analysis based upon our critical knowledge of the subject [43, 44]. The results are therefore interpretable on a human scale. Sentiment analysis, particularly when measured over a narrative of texts, conversely reveals

patterns of affect that are latent within a narrative: by actively reducing the noise of the analysis, we can generalize text, and even corpus-level effects from a broad range of specific, extremely localized, phenomena [45].

Digital analyses conducted for this work use several textual analysis and natural language processing packages. Texts are cleaned with manual removal of paratextual information like bibliographic data and indices, then via Python and R scripts to lowercase the text and remove nontextual information like line breaks. Words are not stemmed, as trials suggest duplicate words do not present a problem for these corpora and because subtle differences in terms like “environment” and “environmental,” “mining” and “miner,” and “gas” and “gasoline” are content-bearing. Stopwords are not removed because of an interest in detecting differences between fiction and nonfiction corpora that are expected to use language differently, although some words are thinned through the topic modeling software that we adopted. Topic modeling is performed using LDA in MALLET (MACHINE Learning for Language Toolkit) [46] (see section S1.1 in the supplementary material for details). Sentiment analysis is performed both using Jockers’ Syuzhet package with the built-in Syuzhet, Bing, NRC, and AFINN dictionaries in R [47] and using VADER (Valence Aware Dictionary for sEntiment Reasoning) in Python [48]. That is, five different dictionary-based sentiment analysis tools are applied to the corpus (see section S1.2), after which the Syuzhet package is used to investigate the emotional arc of each text. Finally, texts are investigated using the NRC emotion lexicon, which identifies words as emotion-associated rather than simply as positive or negative [49].

3 Results

3.1 Topic modeling for latent assumptions

This work presents insights from topic models based on two particular areas of interest: the nature of a text’s portrayal of either energy resource (as a character versus as setting) and the implied relationship of either resource to the economy. However, many other focal points and interpretations are potentially worth investigation. We have elected to make the raw output data from our topic modeling available as supplementary material to other researchers to explore both the logic of our interpretation and other potential findings from our analysis. The supplemental results (section S2.1) includes the specific topics supporting claims made in this main text. Though a strength of topic modeling is that the algorithm naively identifies groups of words that are likely to appear together and can be cautiously interpreted as “themes” by a human reader [50], topics do require subjective interpretation.

3.1.1 Topic modeling the entire corpus

Intuitively, one would expect that the topic models for small numbers of topics k would reflect the basic contours of the corpus as comprising oil and coal fiction and nonfiction texts, as well as reflect the differences between individual texts and sub-corpora. The model performs as expected, with energy- and non-energy topics that split fiction and nonfiction texts emerging for $k = [2,3]$ (full corpus). This result, which also includes topics that reflect oil’s relationship with global trade and corporations and coal’s relationship with American labor ($k = 3$, full corpus), provides confidence that the model is identifying real patterns of association in the texts.

Modeling many more topics than texts (e.g., $k = 300$ for this 60-text corpus) enables high resolution analysis of themes in the overall corpus by making it more likely that less obvious

patterns become visible. That is, while a $k = 2$ model effectively provides fictional and nonfictional topics, a $k = 300$ model offers an opportunity to identify specific issues that are present in the corpus. For example, the environment, mountaintop removal, and labor unions are all visible, as is the BP oil spill. Similarly, people, human interactions, and family are present across fiction texts. Topics that are well distributed across all four major corpus attributes—oil, coal, fiction, and nonfiction—are uncommon: examples include topics with high frequency words that address time, location, and work. Minor topics to highlight include a coal topic that addresses the spiritual and supernatural and an oil topic that deals with money (see section S2.1 for support here and for other topic modeling-based claims).

3.1.2 Topic modeling the nonfiction and fiction subcorpora

Topic models of the nonfiction and fiction subcorpora are the most revealing of all the topic models created for this work with respect to major questions of interest related to the relationship between oil, coal, and the economy and the nature of oil and coal as characters or elements of a narrative's setting. In each case, large k topic models ($k = 5n$, $n =$ number of texts) are used to ensure that rare topics are identified. Though most of the topics generated in the large k model are associated with a single text, identifying groups of topics that look similar to a human reader can successfully provide insight at higher resolution than small k models that tend to group primarily based on narrative genre, plot (e.g. the Deepwater Horizon disaster addressed in multiple oil nonfiction texts), and author style.

3.1.2.1 Topic modeling the nonfiction subcorpus

A high resolution topic model of the nonfiction subcorpus suggests that oil and Appalachian coal have very different relationships to the economy. Specifically, oil brings wealth while coal does not. This relationship is evaluated based on a close examination of topics including both resource-related words (e.g. coal, oil, mine, shale) and economy-related words (e.g. econ*, money). Oil-related economy topics tend to include references to international markets, wealth, and large-scale success through signifiers like the word “billion”. By contrast, coal-related economy topics appear to reference past, lost success (for example, when “economic” appears with “sacrifice” and “heritage”) and activism, suggesting that the economy is appearing in the context of economic development needs.

3.1.2.2 Topic modeling the fiction subcorpus

A high resolution topic model of fiction reveals how oil and coal are treated in fictional texts—whether more like characters or as setting. The result is striking. For the $k = 150$ topic model, about a quarter prominently include multiple characters' names, which is typical for a fictional corpus this size). However, there is a strong correlation between resources and characters, as “oil” and “coal” overwhelmingly appear in the topics that feature character names. Of the 11 topics that include “oil” as a top-20 word, ten include at least four character names in the top 20 words (and nine include at least six). The single exception is a topic where oil is associated with terms related to work and happiness, like “laughed,” “smiled,” “love” and “business,” “company,” and “partner.” Similarly, “coal” appears in six topics and “mine” in nine topics, with 12 of the 15 appearances in character topics. “Coal” and “mine” appear with near-character terms like “brother,” “uncle,” and “father.” “Mine” also appears as “mined”, where it appears to be acting as setting (as in “the area was mined”). This result strongly suggests that both oil and coal are frequently used in more character-like than setting-like contexts.

3.2 Sentiment and the emotional arcs of oil and Appalachian coal narratives

Sentiment in each text is further based on measurement at 2% intervals (Syuzhet dictionary results: Figures 2, 3, 4; Bing, NRC, AFINN, and Vader dictionary results, Figures S1-S4). Figure 1 shows that each subcorpus is, on average, essentially neutral. Sentiment in coal texts is generally less positive than in oil texts, and nonfiction texts for both resources have higher average sentiment and a wider sentiment range than fiction texts (Table 1, Syuzhet dictionary). These observations are stable across all five dictionaries used in this analysis (Tables S2-S5).

Figure 1. Mean sentiment (Syuzhet dictionary) by subcorpus is plotted on the same scale as emotional valence for individual texts (Figures S5-S8) to clearly show the near-0 average sentiment

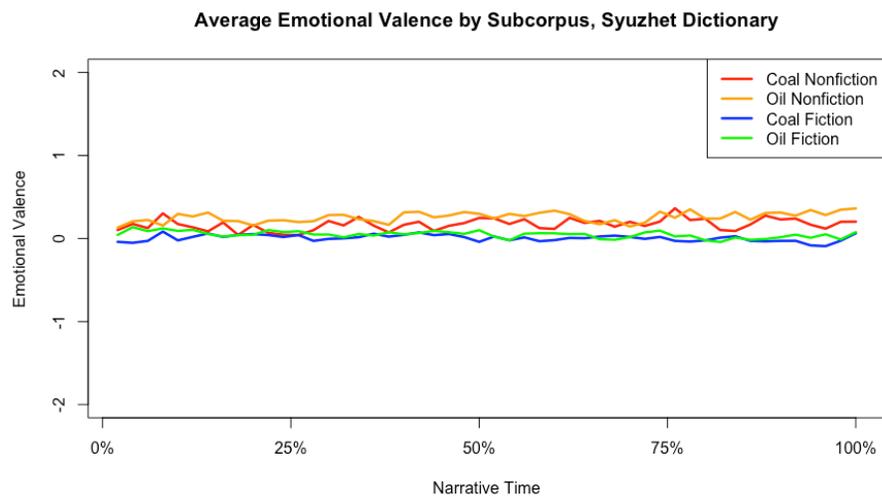


Table 1. Summary Sentiment Analysis Statistics, Syuzhet Dictionary

	Coal fiction	Coal non-fiction	Oil fiction	Oil non-fiction	Coal	Oil	Fiction	Non-fiction
Average sentiment	0.01	0.19	0.05	0.26	0.09	0.15	0.03	0.23
First half average	0.02	0.16	0.07	0.24	0.09	0.16	0.04	0.20
Second half average	-0.01	0.22	0.03	0.28	0.10	0.15	0.01	0.25
Minimum	-0.54	-0.74	-0.44	-0.49	-0.74	-0.49	-0.54	-0.74
Maximum	0.43	1.69	0.49	1.94	1.69	1.94	0.49	1.94

Figures 2 and 3 show high resolution emotional arcs for each text, displayed as heat maps and organized by year of publication. Two years are highlighted to reflect important and widely reported events for US oil and Appalachian coal. In 2008, both oil and US coal experienced unusually high commodity prices. Both resource industries experienced major fatal disasters in April 2010, with the Upper Big Branch Appalachian coal mining disaster and the offshore Deepwater Horizon BP oil spill.

Figure 2. Fiction texts. Darker values of green indicate more positive sentiment, while darker values of red indicate more negative sentiment.

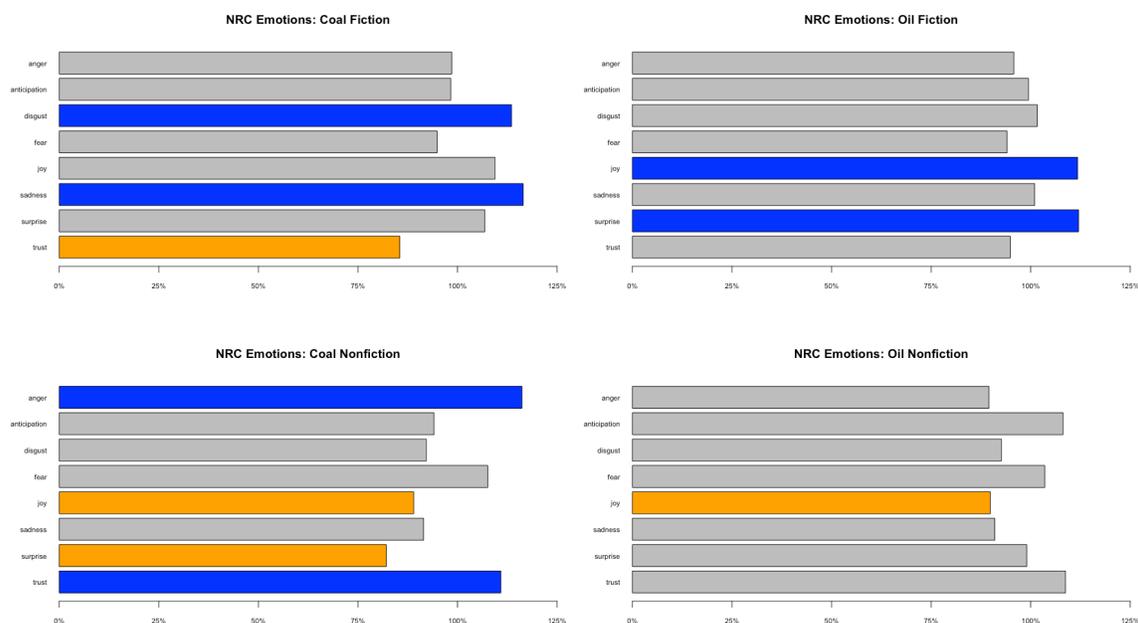


Figure 3. Nonfiction texts. Darker values of green indicate more positive sentiment, while darker values of red indicate more negative sentiment.



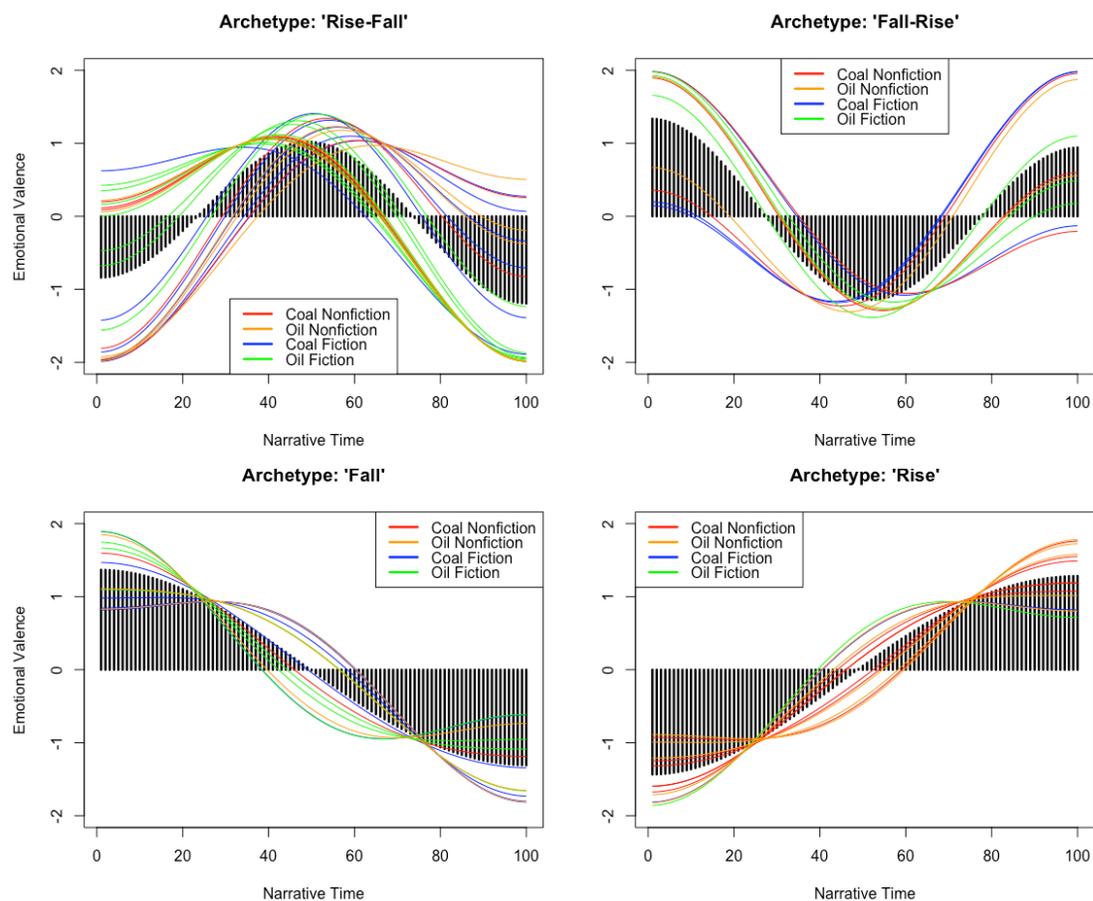
In addition to the evaluations of positive and negative sentiment in the corpus shown in Figures 1-3 and S1-S8, the NRC dictionary enables assessment of a corpus' use of Plutchik's eight basic emotions via a human-annotated emotion dictionary of about 14,000 terms [49]. Figure 4 shows how the emotions are represented in each subcorpus relative to the overall corpus, with emotions either over- or underrepresented by at least 10% highlighted in blue or orange respectively. Negative emotions tend to be highly overrepresented in coal narratives, while positive words tend to be overrepresented in oil narratives.

Figure 4. Emotional appearance in each subcorpus relative to the overall corpus (blue bars show emotions that are overrepresented by at least 10%; orange bars show emotions that are underrepresented by at least 10%)



Another approach to measuring emotion in a text is to evaluate a simplified emotional arc, sometimes considered a proxy for basic plot (see e.g. [51]). Based on results from Jockers' Syuzhet package ([47], and see S1.2 in the supplementary materials), we observe four primary emotional arcs in the corpus of oil and Appalachian coal narratives. Figure 5 presents the emotional arcs of texts associated with four basic patterns, based on the Syuzhet sentiment dictionary: rise-fall, fall-rise, fall, and rise, using the simple descriptors of [52]. Distribution of texts among the observed archetypes is presented in Table S6.

Figure 5. Emotional arcs in the corpus (Syuzhet implementation of Syuzhet dictionary, discrete cosine transformation)



4 Discussion

Sentiment analysis, topic modeling, and other techniques support the hypothesis that latent cultural assumptions are much more positive for oil than Appalachian coal in American narratives. Both resources are most often used in character-like contexts, often given agency and moral value. Oil represents a sometimes fickle alternative reality, but one that is valiant, with the potential for success and wealth. Coal, by contrast, represents an undeniable reality of villainous failure to deliver prosperity or even meet basic expectations of the many human characters that remain loyal to it.

Our theory is confirmed by use of parts of speech in the corpus. It is true that both coal fiction and nonfiction have a lower percentage of past tense verbs (compared to all verbs) than oil fiction and oil nonfiction respectively (coal vs. oil fiction at 0.51 vs. 0.52, and coal vs. oil nonfiction at 0.45 vs. 0.51). This is put into context, however, by the much higher percentage of modal verbs (compared to all verbs) in narratives about oil (both fiction and nonfiction). Modal verbs (will, could, would, should, etc.) are speculative terms, whether projecting a potential future (will, could) or an alternate present (would, should). Oil narratives contain a significantly higher percentage of these words (0.056 and 0.060 for fiction and nonfiction respectively) when compared to coal narratives (0.032 and 0.050 for fiction and nonfiction respectively). This

suggests that oil narratives are indeed more future-looking and speculative, as compared to the reality of coal's present and past.

4.1 Resources as Characters

Oil and coal both appear to be treated more as characters than as elements of setting in the fictional works of this corpus, supported by topic modeling results described in Section 3.1.2.2. This observation is consistent with LeMenager's observation of oil through Sinclair's *Oil!* as an "excessively embodied figure" that is "apparently more alive than its human witnesses" [53]. Dispersion plots (section S1.3 and Figure S9) show that the words "coal" and "oil" and their most common collocates are referenced both frequently and consistently over the course of their respective fictional narratives, also supporting the idea that the resources are present and important to the plot. Detection of emotional arcs in the nonfiction texts that are similar to those found in fictional narratives (Figures S5-S8) further suggests both that stylistically, these books present narrative history, and that the resources have character-like features, with many nonfiction texts effectively presenting a 'biography' of a resource-related event.

These digital results correspond well with observations based on manual review that both resources are given identity and agency, often as antagonists for the human characters. In *Strange as this Weather has Been* [54], Ann Pancake depicts coal as a beloved murderer in a scene describing a man dying from exposure to coal dust in the mines:

His lungs are being buried by it, by coal, which is earth, which is this place, and still, he wants nothing but to be out in it. On the land, like me, like us, despite the burying it does, and what the hell? What the hell is it? Why do we have to love it like we do.

Similarly, Parker Avriale opens *The Runaway Millions* [55] with the dramatic bankruptcy of an extremely wealthy man as a drop in oil price forces him into default:

‘How can you pop up out of nowhere and take everything just because the price of oil dipped below fifty dollars a barrel for a few days? We all know it’s going back up.’...
‘We’re stealing nothing. You gambled, and you lost.’

Coal causes death and oil causes loss of a lifestyle. In these and other texts, both resources repeatedly create major conflict. Frequently, it is the resource itself that is identified as the proximate cause of a trouble, not a human, though humans might take advantage of a situation caused by the antagonizing resource.

In some cases, coal and oil are used more indirectly to denote antagonism, usually by signaling that some person or organization is going to be trouble. For example, fossil fuel companies and their associates are often depicted in negative lights, and the idea of foreignness is also widely deployed to indicate something negative. Grisham's *Gray Mountain* [56] references a coal company explicitly identified as Russian—a relatively common shorthand for a suspicious organization (see e.g. [57, 58])—even though coal is almost exclusively a domestic resource managed by domestic companies. In Margolin's *Violent Crimes* [59], while the oil company is moderately important to the plot, an early mention of a coal company called Global Mining seems explicitly intentioned to establish that the law firm that serves it is not trustworthy. Here

again, the idea of a “global” coal mining company is present despite the reality of primarily domestic operators. Lyn’s *The Sands of Santa Rosa* [60] also sets up the opposition of fossil fuel corporation versus environmental nonprofit, though she inverts a common pattern by deploying a disgruntled environmentalist against an oil company to cause a major spill (ultimately, the spill is disastrous because of steps taken by the oil company, however). While oil and coal are not used as characters in their own rights in these contexts, they provide crucial characterization, doing important narrative work of establishing how readers should interpret characters. The more consistent use of oil and coal to describe characters rather than places further supports the notion that the resources are more character-like than setting-like.

4.2 The Role of Women

Oil and coal are often depicted as antagonizing characters in the fictional corpus analyzed here. Who, then, are the protagonists that stand opposite these resources? Conspicuously, women frequently appear as environmentally and socially oriented characters constructed to create tension with the destructive character-resources. This usage reflects a long history of cultural perception of women as “closer to nature” [61], based on assumptions that women are more connected to and have a special knowledge of the natural world that enables women to more effectively protect the environment and nurture communities (see e.g. [62, 63]). (It is worth noting that while oil and coal are themselves natural resources, their association with technology, industrialization, and environmental harm somewhat erases association to “nature.”) Many have noted either real or perceived differences in participation in environmental work and environmentally friendly behaviors among women, whether mediated through culture or some more essentialist source (see e.g. [63-67]). Other scholars have described the sometimes harmful effect of this stereotype about women as caretakers, noting issues like the tendency to assume that women will willingly assume (additional) unpaid labor to fulfil duties to the environment (see e.g. [61, 68]). Similar stereotypes about indigenous Americans as possessing a privileged relationship to nature (see e.g. [69]) are also relatively common (and visible in this corpus in Lyn’s 2013 novel [60], which includes a part-indigenous protagonist who saves the day with his mystical “Sight”). In this corpus, however, the use of women characters to indicate alignment with nature and a caretaking ethic is most visible and is the focus of this discussion.

The fossil fuel-oriented contemporary American corpus assessed here uses stereotypes about women as cultural shorthands: they generally offer a moral and eco-connected contrast to the fairly flat depictions of immoral and destructive fossil fuels. Texts like Dorgan and Hagberg’s *Blowout* [70], Grisham’s *Gray Mountain* [56], and Margolin’s *Violent Crimes* [59] are examples of novels by male authors featuring female protagonists battling coal and oil for the good of their community, society, or humanity. In all three cases, the women are excessively accomplished attractive, go-getting truth tellers whose major roles involve opposing conventional fossil fuel development in some form. Former Senator Dorgan and novelist Hagberg [70] feature two women, an implausibly brilliant scientist working against coal’s climate impacts and a scrappy journalist working to reveal the truth about the dangerous project to the public. Grisham’s *Gray Mountain* [56] centers on a nonprofit legal organization entirely staffed by women helping the poor in Appalachian coal country—some of whom need help because of complications from coal mining. Margolin’s *Violent Crimes* [59] pits a female protagonist against a major law firm that serves both coal and oil companies. While Margolin often writes female leads, Grisham’s and Hagberg’s protagonists are more commonly men, suggesting that the choice of women as

protagonists in novels opposing fossil fuels is meaningful. We note this finding is particularly evident in novels by male authors. At least within this corpus, women who write women are less likely to characterize their protagonists as overtly pro-social representatives of a fossil-opposing class (see, for example, [71, 72]).

A striking point of support for the hypothesis that, in this corpus, women are perceived as necessary to highlight pro-social and pro-environmental activities and viewpoints in fossil fuel narratives is the fact that the only woman protagonist in the oil nonfiction corpus is fictional. Jessica Pherma is a character created as an audience surrogate to advance the plot in Turley's *The Simple Truth: BP's Macondo Blowout* [73]. Pherma is described in the text as someone who "would have raised a flag and declared 'STOP WORK!'" [73] to save the day during the Macondo disaster. This characterization both supports the notion that women are seen as caretakers and environmental stewards and echoes the use of pro-social female protagonists in fiction.

Elsewhere in the nonfiction corpus, we find that contemporary coal nonfiction explicitly addresses women both generally and specifically, but contemporary oil nonfiction largely does not. This assessment is supported by topic modeling. The assumed and described role of women as pro-environment and anti-establishment, in contrast with men, is visible in a $k = 3$ topic model of the coal nonfiction corpus, for example. One topic relates "women" to "environmental," "community," "local," "justice," "social," and "people" in its top 20 most characteristic words, while a second topic relates "men" to "miners," "workers," "union," "labor," and "company."

Perhaps notably, neither "women" nor "men" appears in a parallel model of the oil nonfiction corpus, though several high-profile men's names do appear. Topic models designed for higher resolution (larger number of topics k) confirm frequent reference to specific men, like Aubrey McClendon, George Mitchell, Henry Kissinger, H. L. Hunt, and John Rockefeller in the oil nonfiction corpus. By contrast, multiple works in the coal nonfiction corpus—far more oriented around environmental justice than the oil nonfiction corpus—reference specific women like Judy Bonds and Maria Gunnoe, leaders in the Appalachian mountaintop removal opposition community, by name. Men, including controversial coal company CEO Don Blankenship and local journalist Ken Ward, also appear. Whether the discrepancy is due to an actual difference in women's prominence in the US oil and Appalachian coal communities, to different choices about the types of events nonfiction writers address related to each resource, to a bias in the corpus, or to something else is unclear and worthy of further investigation. Future work might productively investigate the role of women in traditionally male fossil fuel industries (building on work like [74]), particularly focusing on whether the often moralistic portrayal of women as environmentally and socially conscious—often at the implicit or explicit expense of making money—is reflected in reality and why.

4.3 Resource Relationship with the Environment, Society, and Economy

Environment, society, and the economy are the major spheres typically associated with sustainability. Given a partly policy-oriented interest in understanding latent cultural attitudes about oil and Appalachian coal in the US, we assess the relationship between each resource and these elements of sustainability in our 60-narrative corpus. Overall, we find that oil is a valiant provider of opportunity, even though it is technical and unfamiliar, sometimes bearing the risk of

pain and death. Negative outcomes are forgiven, and the promise of future economic prosperity is foregrounded. By contrast, coal is a villain that is deeply integrated into the social and environmental fabric of a place but fails to bring economic prosperity. Negative outcomes are betrayals, and people treasure their coal-associated histories even as they recognize the many negative effects of coal. Conlogue writes of “the love of a remembered landscape” and the poet Merwin’s realization that Americans—in this context, Appalachian Americans—are surprised at discovering that “the name for what you feel is love” for a landscape [28].

In general, the coal corpus engages socially and locally oriented issues more directly than the oil corpus, including discussion of environmental justice. By contrast, the oil corpus separates environmental issues from affected communities, often using more technical language. Words commonly co-located with the terms “coal” and “oil” in their respective corpora reveal this distinction. The technical terms co-located with coal are at human scale: words like “processing,” “seam,” “pillars,” and “coking.” Other common co-locations include other human and personal terms like “chunks,” “load,” “stove,” and “miner.” By contrast, “oil” is co-located with exotic and financial terms like “Iranian,” “foreign,” “futures,” “derivative,” “consortium,” and “wealth.” At higher resolution, consider what topic models reveal about latent associations with two issues seen as environmental disasters with large impacts on local communities: the BP oil spill and mountaintop removal coal mining. While a topic related to the BP spill includes technical oil well terms like “bop” (for blowout preventer, a piece of equipment that failed during that event), “riser,” “rams,” “incident,” and “dispersant”, a parallel mountaintop removal mining topic is far more socially oriented, including terms like “land,” “nature,” “white,” “racial,” and “identity.”

The exception to the trend for coal texts to be more locally oriented than oil texts is visible as an environmental justice topic in the oil subcorpus, which draws heavily on Ottinger’s *Refining Expertise* [75] about a community near a refinery. The nature of this text suggests a potential reason for the difference in environmental scale described in the oil and coal subcorpora: it focuses on a single community near a single facility. Appalachian coal is typically extracted and processed at a few large, long-lived facilities (mines) in a specific part of the United States, while US oil is typically extracted at many small facilities (wells) in communities that cannot easily be summarized. Further, much of the US’ oil is produced in other countries or in remote areas like Alaska, offshore Gulf of Mexico, and sparsely populated parts of Texas. As large, long-lived facilities, refineries are more like coal mines than oil wells in a culturally important way. This distinction might explain the more substantial environmental justice focus in refining versus drilling contexts.

The appearance of environmental justice in the corpus indicates that environmental and social issues are often entwined. The tension between communities and development of energy resources that entail large-scale changes in the human environment—where people live, work, and raise families—is a persistent theme in research (e.g. [21, 76-78]) and is also evident in the nonfiction analyzed here. For example, in one topic model of the nonfiction texts ($k = 6$), only one topic meaningfully draws from both oil and coal texts. This single blended topic addresses the relationship between resources and society, with “industry” and “community” as its top two words. The emergence of this topic suggests that the industry-community dynamic is perhaps the biggest shared issue for the two resources. Despite differences in latent assumptions about oil

and coal, and despite very different framings of the impact that either industry has on a place, development of both resources is affected by industry relationships to human communities.

Industry/community relationships are frequently tense, as reflected in one of the corpus texts' titles: *Combating Mountaintop Removal: New Directions in the Fight against Big Coal* [79]. However, a closer reading shows that industry/community relationships are often discussed in ways that imply effort towards (if not always success in) collaborative, low-conflict interactions, as Gwen Ottinger describes in *Refining Expertise* [75]:

With regulators, industry officials, and even local newspaper editors pushing the community toward talks, CCNS's invitation to Clark Johnson might have been little more than a strategic gesture—a superficial repackaging of campaign demands in the locally popular language of discussion. Yet CCNS members were actually quite sincere: they longed to be able to settle their differences with Orion through less conflictual means than marches and lawsuits. Their ideal, as expressed by CCNS leader Don Winston, was to be able to sit down with Orion's top decision makers and talk “like reasonable businesspeople.”

Another aspect of the industry/community relationship regards economic outcomes associated with the oil and coal industries. “Valiant” oil is often associated with grand effects on the economy—wealth and money are described with words like “billions” and “global” (Section 3.1.2.1). As Buell writes, oil is culturally characterized by outsized experiences of “exuberance and catastrophe” [9]. Surprise and anticipation are emotionally overrepresented in the oil subcorpus (Figure 4). Close reading confirms that oil is often associated with gambling, where wins and losses are large, exciting, and ultimately out of an individual's control (see, for example, Jiles' *Stormy Weather* [55] and Avriole's *The Runaway Millions* [72]). With oil, booms and busts occur quickly enough that a person in a bust period can hope for the next boom soon. And when booms are good, they create lasting wealth that feeds entire subcultures (as in DiSclafani's *The After Party* [71]).

In contrast to the exciting, rapid, and ultimately individual-focused success and failure periods observed with oil (see also [53], chapter 2 on the spectacular nature of oil and the illusion of the resource as democratic), “villainous” coal provides entire communities with wealth for a relatively long period of time, then disappears, leaving communities with a legacy of poverty. Appalachia's particular history of labor disputes and relatively thin “boom” times—when miners made low wages and died of black lung even in good period—makes reflections on these good times even starker in comparison to economic distress reflected in the present of many narratives in this corpus. The economy is referenced in terms of economic distress, economic development opportunities, and other phrases that suggest difficulty, tough times, and lengthy periods of decline rather than expectations for overnight fortunes seen in many of the oil texts. Money is associated with words like “need,” “hard,” and “trouble” rather than the “billion,” “love,” and “larger” of the oil corpus. Dominant emotions in the coal corpus are disgust and sadness (in fiction); anger and a lack of joy or surprise (in nonfiction) (Figure 4). Notably, trust is overrepresented in coal nonfiction but underrepresented in coal fiction. One explanation lies in the multifaceted valence of “trust”: nonfiction works tend to reflect a history of cultural constancy and familiarity (associated with trust, even if it is a trust that negative things will

happen), while fiction reflects the experience of an individual or community whose trust is betrayed.

Oil is a resource of hope and excitement, while coal is a resource of disappointment and hardship. Even though both resources have generations of history in the United States, oil remains a potential future while coal is a challenging past. Even modern-day depictions of coal tend to focus on the future past that the present creates: mountaintop removal mining is portrayed as a way to rob the future, in contrast to oil's perceived ability to provide a future. For an example from nonfiction, contrast Joyce Barry's description of mountaintop removal coal mining in *Standing Our Ground: Women, Environmental Justice, and the Fight to End Mountaintop Removal* [80]:

...I was shockingly disturbed that [mountaintop removal] was legal and occurring in my home state...Coming of age in West Virginia, the beautiful mountains that surrounded us were inextricably linked to our history, culture, and sense of place in the world. To learn they are now razed for coal extraction, **left in ruins by heavy machinery and the technicians who operate them**, is simply unacceptable... (emphasis added)

with Russell Gold's description of hydraulic fracturing for oil and natural gas in *The Boom* [81]:

To a remarkable extent, this once-obscure oil-field technique **defines the nation's economic and environmental future**. Fracking has unleashed more oil and natural gas than anyone thought possible. It **is providing** an abundance of domestic energy, helping to drive a rebirth of manufacturing, and easing dependence on overseas energy peddlers. ...It also requires turning whole counties into industrial zones, complete with fleets of trucks, air quality concerns, a disruption of nature, and fear that water aquifers **will be poisoned**. (emphasis added)

Both mountaintop removal and hydraulic fracturing are relatively recently deployed techniques for fossil fuel extraction that have attracted intense public attention, particularly related to environmental degradation and social disruption. Barry writes of mountaintop removal as a future past, with mountains left in ruins. Gold, by contrast—even though he also uses harsh terms to describe environmental challenges—remains future-oriented and somewhat hypothetical throughout.

In a similar comparison from the fiction corpus, Ann Pancake's *Strange as this Weather Has Been* [54] portrays a coal-related disaster that brings death:

It said the impoundment had been sitting on top of a mountain just honeycombed with abandoned deep mines, and finally its bottom simply gave out. The sludge lake dropped into the shafts, shot through the mine tunnels and out through the blocked-up drift mouths, and 306 million gallons of poison muck killed everything in the waterways for a hundred miles.

Meanwhile, an oil-related disaster written (though not set) around the same time, in Paulette Jiles' *Stormy Weather* [72], is mitigated by an almost joyful expectation of future prosperity:

Men ran for their lives. It cost a crew ten days' labor to shut it down. They had hit the biggest oil pool in the history of the world and it was sweet, high-gravity oil so pure you could almost pour it straight into your gas tank, it was the color of honey.

In both cases, the disaster is described in detail, with rich imagery and quantitative description. But oil merits forgiveness by bringing opportunity and prosperity, while coal is a long-term destroyer that will only bring more death and destruction.

4.4 Analytical Challenges and Sources of Error

While digital methods provide insights that are difficult to access as a human reader [38], they pose several challenges that should be noted. In particular, digital sentiment analysis remains a challenge, not least because more nuanced views of sentiment than “positive” and “negative” are still unusual (notably, human analysts often have difficulty replicating others' assessment of sentiment beyond positive and negative, so this is likely to remain a problem long term). While sentiment lexicons and classification rules have improved at identifying tricky constructions like “not good” versus “good,” work to improve these methods continues. For this particular work, it should also be reiterated that the corpus is nonrandom and might not be representative of contemporary works by American authors that address oil and Appalachian coal.

5 Conclusions

A digital investigation of American energy narratives, focused on recent depictions of oil and Appalachian coal in fictional and nonfictional literature, provides insights on the latent cultural attitudes surrounding these two important resources. As policy actions to address issues like climate change and other forms of energy pollution or social impacts continue to develop, remaining aware of the cultural context of resources is likely to be relevant (see e.g. [11]). Whether consciously or subconsciously, oil is associated with more positive outcomes like wealth and prosperity than coal, which might be instructive in considering public reaction to policy proposals. Further, recognizing the contribution of long-term histories for each resource to modern attitudes reveals interesting differences in treatment of the two resources. Despite the fact that in the last several years, each resource has been associated with a major fatal accident, dramatic price swings, politically charged boom and bust cycling, and major shifts in extraction methods that increase access to the resource but pose major environmental challenges, the depiction of these events is very different. Understanding the context of oil's portrayal as a personal gamble where an individual can create legendary personal wealth versus coal's portrayal as a bitter fight between employer and employed helps to situate this observation. Modern American literature depicts oil as the valiant bringer of a potential, exciting alternative reality, while coal is the villain of a familiar, disappointing history.

6 Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. EG is supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-114747. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

7 References

- [1] Gopalakrishnan, Sathya, and H. Allen Klaiber. 2014. "Is the Shale Energy Boom a Bust for Nearby Residents? Evidence from Housing Values in Pennsylvania." *American Journal of Agricultural Economics* 96 (1): 43–66. doi:10.1093/ajae/aat065.
- [2] Jacquet, Jeffrey B., and Richard C. Stedman. 2014. "The Risk of Social-Psychological Disruption as an Impact of Energy Development and Environmental Change." *Journal of Environmental Planning and Management* 57 (9): 1285–1304. doi:10.1080/09640568.2013.820174.
- [3] Tabassum-Abbasi, M. Premalatha, Tasneem Abbasi, and S. A. Abbasi. 2014. "Wind Energy: Increasing Deployment, Rising Environmental Concerns." *Renewable and Sustainable Energy Reviews* 31 (March): 270–88. doi:10.1016/j.rser.2013.11.019.
- [4] Willow, Anna, and Sara Wylie. 2014. "Politics, Ecology, and the New Anthropology of Energy: Exploring the Emerging Frontiers of Hydraulic Fracking." *Journal of Political Ecology* 21 (12): 222–236.
- [5] Bruner, Jerome. 1991. "The Narrative Construction of Reality." *Critical Inquiry* 18 (1): 1–21.
- [6] Geertz, Clifford. 1973. "Thick Description: Toward an Interpretive Theory of Culture." *Culture: Critical Concepts in Sociology* 1: 173–196.
- [7] Oppermann, Serpil. 2006. "Theorizing Ecocriticism: Toward a Postmodern Ecocriticism Practice." <http://www.academia.edu/download/943055/Opperman.pdf>.
- [8] Glotfelty, Cheryll. 1996. "Introduction: Literary Studies in an Age of Environmental Crisis." In *The Ecocriticism Reader: Landmarks in Literary Ecology*, edited by Cheryll Glotfelty and Harold Fromm, xv–xxxvii. University of Georgia Press.
- [9] Buell, Frederick. 2012. "A Short History of Oil Cultures: Or, the Marriage of Catastrophe and Exuberance." *Journal of American Studies* 46 (2): 273–93. doi:10.1017/S0021875812000102.
- [10] Kneitz, Agnes. 2015. "'As If the River Was Not Meat and Drink to You!': Social Novels as a Means of Framing Nineteenth-Century Environmental Justice." *Interdisciplinary Studies in Literature and Environment* 22 (1): 47–62. doi:10.1093/isle/islv006.
- [11] Spurgeon, S. L. 2009. "Miracles in the Desert: Literature, Water, and Public Discourse in the American West." *Interdisciplinary Studies in Literature and Environment* 16 (4): 743–59. doi:10.1093/isle/isp066.
- [12] EIA. 2016. "Shale in the United States." http://www.eia.gov/energy_in_brief/article/shale_in_the_united_states.cfm.
- [13] Yergin, Daniel. 2011. *The Quest: Energy, Security, and the Remaking of the Modern World*. Upd Rev Re edition. Penguin Books.
- [14] Zuckerman, Gregory. 2013. *The Frackers: The Outrageous Inside Story of the New Billionaire Wildcatters*. Reprint edition. Portfolio.
- [15] Houser, Trevor, and Shashank Mohan. 2014. *Fueling Up: The Economic Implications of America's Oil and Gas Boom*. Peterson Institute for International Economics.
- [16] Boudet, Hilary, Christopher Clarke, Dylan Bugden, Edward Maibach, Connie Roser-Renouf, and Anthony Leiserowitz. 2014. "'Fracking' Controversy and Communication: Using National Survey Data to Understand Public Perceptions of Hydraulic Fracturing." *Energy Policy* 65 (February): 57–67. doi:10.1016/j.enpol.2013.10.017.
- [17] Safford, Thomas G., Jessica D. Ulrich, and Lawrence C. Hamilton. 2012. "Public Perceptions of the Response to the Deepwater Horizon Oil Spill: Personal Experiences,

- Information Sources, and Social Context.” *Journal of Environmental Management* 113 (December): 31–39. doi:10.1016/j.jenvman.2012.08.022.
- [18] Starbird, Kate, Dharma Dailey, Ann Hayward Walker, Thomas M. Leschine, Robert Pavia, and Ann Bostrom. 2015. “Social Media, Public Participation, and the 2010 BP Deepwater Horizon Oil Spill.” *Human and Ecological Risk Assessment: An International Journal* 21 (3): 605–30. doi:10.1080/10807039.2014.947866.
- [19] EIA. 2016. “Short-Term Energy Outlook.” <https://www.eia.gov/forecasts/steo/realprices/>.
- [20] Schafft, Kai A., Yetkin Borlu, and Leland Glenna. 2013. “The Relationship between Marcellus Shale Gas Development in Pennsylvania and Local Perceptions of Risk and Opportunity.” *Rural Sociology* 78 (2): 143–66. doi:10.1111/ruso.12004.
- [21] Ellis, Colter, Gene Theodori, Peggy Petrzela, Douglas Jackson-Smith, and A.E. Luloff. 2016. “Unconventional Risks: The Experience of Acute Energy Development in the Eagle Ford Shale.” *Energy Research & Social Science*, June. doi:10.1016/j.erss.2016.05.006.
- [22] Grattan, Lynn M., Sparkle Roberts, William T. Mahan, Patrick K. McLaughlin, W. Steven Otwell, and J. Glenn Morris. 2011. “The Early Psychological Impacts of the Deepwater Horizon Oil Spill on Florida and Alabama Communities.” *Environmental Health Perspectives* 119 (6): 838–43.
- [23] Morris, J. Glenn, Lynn M. Grattan, Brian M. Mayer, and Jason K. Blackburn. 2013. “Psychological Responses and Resilience of People and Communities Impacted by The Deepwater Horizon Oil Spill.” *Transactions of the American Clinical and Climatological Association* 124: 191–201.
- [24] Ross, Jonathan Knox. 2009. “Offshore Industry Shift Work—health and Social Considerations.” *Occupational Medicine* 59 (5): 310–15. doi:10.1093/occmed/kqp074.
- [25] Haslam McKenzie, Fiona M., and Aileen Hoath. 2014. “The Socio-Economic Impact of Mine Industry Commuting Labour Force on Source Communities.” *Resources Policy* 42 (December): 45–52. doi:10.1016/j.resourpol.2014.09.002.
- [26] EIA. 2016. “Coal Data Browser.” <http://www.eia.gov/beta/coal/data/browser/>.
- [27] Anglin, Mary K. 2002. “Lessons from Appalachia in the 20th Century: Poverty, Power, and the ‘Grassroots.’” *American Anthropologist* 104 (2): 565–82. doi:10.1525/aa.2002.104.2.565.
- [28] Conlogue, B. 2011. “Merwin and Mining.” *Interdisciplinary Studies in Literature and Environment* 18 (3): 594–614. doi:10.1093/isle/isr078.
- [29] Hirsch, Susan F., and E. Franklin Dukes. 2014. *Mountaintop Mining in Appalachia: Understanding Stakeholders and Change in Environmental Conflict*. Chapter 2. Ohio University Press.
- [30] Lewis, Ronald L. 1993. “Appalachian Restructuring in Historical Perspective: Coal, Culture and Social Change in West Virginia.” *Urban Studies* 30 (2): 299–308. doi:10.1080/00420989320080301.
- [31] Scott, Rebecca R. 2010. *Removing Mountains: Extracting Nature and Identity in the Appalachian Coalfields*. Chapter 5. Univ Of Minnesota Press.
- [32] Absher, Brandon, and Radical Philosophy Association, in cooperation with the Philosophy Documentation Center. 2012. “Toward a Concept of Ecological Violence: Heidegger and Mountain Justice.” Edited by Eduardo Mendieta. *Radical Philosophy Review* 15 (1): 89–101. doi:10.5840/radphilrev20121519.

- [33] Bell, Shannon Elizabeth, and Richard York. 2010. "Community Economic Identity: The Coal Industry and Ideology Construction in West Virginia." *Rural Sociology* 75 (1): 111–43. doi:10.1111/j.1549-0831.2009.00004.x.
- [34] Argamon, Shlomo, Charles Cooney, Russell Horton, Mark Olsen, Sterling Stein, and Robert Voyer. 2009. "Gender, Race, and Nationality in Black Drama, 1950-2006: Mining Differences in Language Use in Authors and Their Characters" 3 (2). <http://www.digitalhumanities.org/dhq/vol/3/2/000043/000043.html>.
- [35] Burrows, John. 2004. "Textual Analysis." In *A Companion to Digital Humanities*, edited by Susan Schreibman, Ray Siemens, and John Unsworth. Oxford: Blackwell. http://digitalhumanities.org/companion/view?docId=blackwell/9781405103213/9781405103213.xml&chunk.id=ss1-4-1&toc.id=0&brand=9781405103213_brand.
- [36] Jockers, Matthew L., and Daniela M. Witten. 2010. "A Comparative Study of Machine Learning Methods for Authorship Attribution." *Literary and Linguistic Computing*, April, fq001. doi:10.1093/lc/fqq001.
- [37] Coll, Ardanuy and C. Sporleder. 2014. "Structure-based Clustering of Novels." *EACL 2014: Proceedings of the 3rd Workshop on Computational Linguistics for Literature*. <http://www.aclweb.org/anthology/W/W14/W14-0905.pdf>
- [38] Grubert, Emily, and Anne Siders. 2016. "Benefits and Applications of Interdisciplinary Digital Tools for Environmental Meta-Reviews and Analyses." *Environmental Research Letters* 11 (9): 93001. doi:10.1088/1748-9326/11/9/093001.
- [39] Jockers, Matthew. 2013. *Macroanalysis: Digital Methods and Literary History*. (U of Illinois P)
- [40] Piper, Andrew. 2016. "There Will Be Numbers." *Journal of Cultural Analytics*. culturalanalytics.org/2016/05/there-will-be-numbers/
- [41] Algee-Hewitt, Mark, Franco Moretti and Ryan Heuser. 2015. "On Paragraphs. Scales, Themes, and Narrative Form." *Pamphlets of the Stanford Literary Lab* (10) litlab.stanford.edu/LiteraryLabPamphlet10.pdf
- [42] Dixon, Dan. 2012. "Analysis Tool or Research Methodology: Is There an Epistemology for Patterns?" In: Berry, D ed. *Understanding Digital Humanities* (Palgrave MacMillian)
- [43] Blei, David M. 2012. "Probabilistic Topic Models *Communications of the ACM* (55.4) doi: 10.1145/2133806.2133826
- [44] Algee-Hewitt, Mark and Bridget F.B. Algee-Hewitt 2016. "The Unspoken Word: Race and the Changing Language of Identity." *Current Anthropology* (forthcoming)
- [45] Jockers, Matthew. 2016. "A multiscale theory for the dynamical evolution of sentiment in novels. 2016 International Conference on Behavioral, Economic and Socio-Cultural Computing doi: 10.1109/besc.2016.7804470
- [46] McCallum, AK. 2002. "MALLET: A Machine Learning for Language Toolkit." <http://mallet.cs.umass.edu>.
- [47] Jockers, Matthew. 2016. "Syuzhet." *GitHub*. Accessed January 30, 2017. <https://github.com/mjockers/syuzhet>.
- [48] Hutto, Clayton J., and Eric Gilbert. 2014. "Vader: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text." In *Eighth International AAAI Conference on Weblogs and Social Media*. <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8109>.
- [49] Mohammad, Saif M., and Peter D. Turney. 2013. "Crowdsourcing a Word–emotion Association Lexicon." *Computational Intelligence* 29 (3): 436–465.

- [50] Blei, David M. 2013. "Topic Modeling and Digital Humanities." *Journal of Digital Humanities*. April 8. <http://journalofdigitalhumanities.org/2-1/topic-modeling-and-digital-humanities-by-david-m-blei/>.
- [51] Jockers, Matthew. 2014. "A Novel Method for Detecting Plot." Accessed January 30, 2017. <http://www.matthewjockers.net/2014/06/05/a-novel-method-for-detecting-plot/>.
- [52] Reagan, Andrew J., Lewis Mitchell, Dilan Kiley, Christopher M. Danforth, and Peter Sheridan Dodds. 2016. "The Emotional Arcs of Stories Are Dominated by Six Basic Shapes." *EPJ Data Science* 5 (1): 31. doi:10.1140/epjds/s13688-016-0093-1.
- [53] LeMenager, Stephanie. *Living Oil: Petroleum Culture in the American Century*. Oxford University Press, 2014. *CrossRef*. Web. 17 Jan. 2017.
- [54] Pancake, Ann. 2007. *Strange as This Weather Has Been: A Novel*. Counterpoint.
- [55] Avrile, Parker. 2016. *The Runaway Millions: A Male/Male Contemporary Romance Novel*. Paris April Press.
- [56] Grisham, John. 2014. *Gray Mountain: A Novel*. Dell.
- [57] Brook, Tom. 2017. "Hollywood Stereotypes: Why Are Russians the Bad Guys?" Accessed January 24. <http://www.bbc.com/culture/story/20141106-why-are-russians-always-bad-guys>.
- [58] Zevin, Jack. 2003. "Perceptions of National Identity: How Adolescents in the United States and Norway View Their Own and Other Countries." *The Social Studies* 94 (5): 227–31. doi:10.1080/00377990309600211.
- [59] Margolin, Phillip. 2016. *Violent Crimes: An Amanda Jaffe Novel*. Harper.
- [60] Lyn, Tommie. 2013. *The Sands of Santa Rosa*. Blackwater Books Publishing.
- [61] Leach, Melissa. 2007. "Earth Mother Myths and Other Ecofeminist Fables: How a Strategic Notion Rose and Fell." *Development and Change* 38 (1): 67–85. doi:10.1111/j.1467-7660.2007.00403.x.
- [62] Gaard, Greta. 2015. "Ecofeminism and Climate Change." *Women's Studies International Forum* 49 (March): 20–33. doi:10.1016/j.wsif.2015.02.004.
- [63] Zelezny, L., and M. Bailey. 2006. "A Call for Women to Lead a Different Environmental Movement." *Organization & Environment* 19 (1): 103–9. doi:10.1177/1086026605285588.
- [64] Bell, Shannon Elizabeth. 2013. *Our Roots Run Deep as Ironweed: Appalachian Women and the Fight for Environmental Justice*. University of Illinois Press.
- [65] Freudenberg, Nicholas, and Carol Steinsapir. 1991. "Not in Our Backyards: The Grassroots Environmental Movement." *Society & Natural Resources* 4 (3): 235–45. doi:10.1080/08941929109380757.
- [66] Prindeville, Diane-Michele, and John G. Bretting. 1998. "Indigenous Women Activists and Political Participation." *Women & Politics* 19 (1): 39–58. doi:10.1300/J014v19n01_03.
- [67] Tindall, David B., Scott Davies, and Celine Mauboules. 2003. "Activism and Conservation Behavior in an Environmental Movement: The Contradictory Effects of Gender." *Society & Natural Resources* 16 (10): 909–932.
- [68] Jackson, Cecile. 1993. "Women/Nature or Gender/History? A Critique of Ecofeminist 'development.'" *The Journal of Peasant Studies* 20 (3): 389–418. doi:10.1080/03066159308438515.
- [69] Krech, Shepard. 2000. *The Ecological Indian: Myth and History*. W. W. Norton & Company.
- [70] Dorgan, Byron L., and David Hagberg. 2013. *Blowout*. New York: Forge Books.
- [71] DiSclafani, Anton. 2016. *The After Party: A Novel*. Riverhead Books.

- [72] Jiles, Paulette. 2009. *Stormy Weather: A Novel*. Reprint edition. HarperCollins e-books.
- [73] Turley, J. A. 2012. *The Simple Truth: BP's Macondo Blowout*. 1 edition. The Brier Patch, LLC.
- [74] Rolston, Jessica Smith. 2014. *Mining Coal and Undermining Gender: Rhythms of Work and Family in the American West*. Rutgers University Press.
- [75] Ottinger, Gwen. 2013. *Refining Expertise: How Responsible Engineers Subvert Environmental Justice Challenges*. NYU Press.
- [76] Anderson, Peggy A. 2014. "Preserving Place: Williston, ND and the Use of the Community Capitals Framework for Boomtown Community Development." *Housing and Society* 41 (2): 297–314. doi:10.1080/08882746.2014.11430632.
- [77] Gasteyer, Stephen, and Jennifer Carrera. 2013. "The Coal-Corn Divide: Colliding Treadmills in Rural Community Energy Development." *Rural Sociology* 78 (3): 290–317. doi:10.1111/ruso.12013.
- [78] Willow, Anna J. 2014. "The New Politics of Environmental Degradation: Un/Expected Landscapes of Disempowerment and Vulnerability." *Journal of Political Ecology* 21 (1): 237–257.
- [79] McNeil, Bryan T. 2011. *Combating Mountaintop Removal: New Directions in the Fight against Big Coal*. University of Illinois Press.
- [80] Barry, Joyce M. 2012. *Standing Our Ground: Women, Environmental Justice, and the Fight to End Mountaintop Removal*. Reprint edition. Ohio University Press.
- [81] Gold, Russell. 2014. *The Boom: How Fracking Ignited the American Energy Revolution and Changed the World*. Reprint edition. Simon & Schuster.