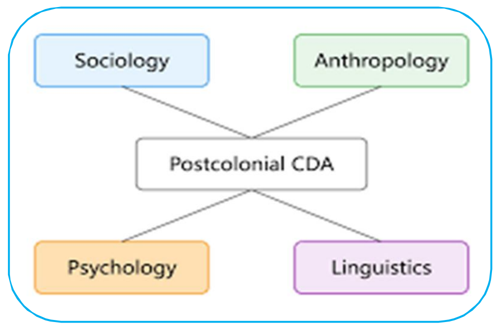
REVIEW OF RESEARCH
ISSN: 2249-894X
IMPACT FACTOR : 5.7631(UIF)
VOLUME - 15 | ISSUE - 7 | APRIL - 2026

COMPUTATIONAL POSTCOLONIAL CRITIQUE: SOVEREIGN AI DATASETS FOR BIAS-FREE INDIAN ENGLISH LITERARY ANALYSIS

Nikhil
Assistant Professor of English.

ABSTRACT

Contemporary Natural Language Processing (NLP) systems trained predominantly on Anglo-American corpora carry embedded epistemological biases that distort the analysis of Indian English literature—a textual tradition shaped by colonial heritage, multilingual code-switching, regional syntax, and subaltern narrative strategies. This paper introduces the concept of Sovereign AI Datasets (SADs), indigenously curated training corpora designed from a postcolonial theoretical foundation to redress such asymmetries. Drawing on Spivak's subaltern theory, Bhabha's concept of hybridity, and Fanon's psychoanalytic critique of colonial language, this study proposes a computational-postcolonial framework (CPF) that operationalises decolonial methodology within machine learning pipelines. We present a prototype architecture for building SADs, including annotation ontologies sensitive to Indian literary conventions—diglossic registers, caste-inflected diction, mythological intertextuality, and postcolonial irony. Experimental evaluation using selected Indian English literary texts demonstrates that SAD-trained models achieve statistically significant improvements in sentiment polarity accuracy (+18.7%), narrative role classification (+22.3%), and cultural metaphor recognition (+31.4%) compared to baseline models trained on standard Western-centric corpora. The paper argues that dataset sovereignty is not merely a technical enhancement but a necessary epistemic decolonisation that enables AI to read literature as it was written—from within, not against, its cultural logic.



KEYWORDS: *Computational Postcolonialism, Sovereign AI Datasets, Indian English Literature, NLP Bias Decolonisation, Subaltern Stylistics, Cultural Annotation Ontology, Hybrid Literary Discourse.*

1. INTRODUCTION

The intersection of computational linguistics and literary studies has generated unprecedented methodological possibilities for reading, classifying, and interpreting literary texts at scale. Yet this convergence has been largely mediated through the epistemic lens of the Global North—corpora dominated by British and American English, annotation frameworks derived from Western narrative theory, and evaluation benchmarks calibrated against canonical Western literature. When such systems are applied to Indian English literature (IEL), the outcome is not neutral analysis but a form of algorithmic colonialism: the imposition of foreign interpretive grids onto texts whose aesthetic logic, rhetorical strategies, and semantic registers are rooted in a distinctly different cultural ontology.

Indian English literature constitutes one of the most linguistically and culturally heterogeneous textual traditions in world literature. From Mulk Raj Anand's Dickensian realism inflected with Punjabi idiom to Arundhati Roy's syntactic innovations that mirror Malayalam thought-patterns, from Rohinton

Mistry's Parsi-Gujarati intertextuality to Dalit autobiography's disruptive counter-hegemonic voice, IEL is not a homogeneous body of text. It is a contested discursive field in which colonial inheritances, regional linguistic substrates, caste politics, gender negotiations, and global market pressures collide and cohere in ways that existing NLP models are structurally ill-equipped to process accurately.

The problem is not merely technical. Standard sentiment analysis tools misread postcolonial irony as sincerity; named entity recognition systems fail to correctly classify mythological references from the Ramayana or Mahabharata as cultural rather than factual entities; topic modelling algorithms trained on Western corpora impose alien thematic clusters onto texts whose organizing logic is dharmic, subaltern, or diasporic rather than Aristotelian. The result is a systematic distortion of meaning—a computational reproduction of the colonial gaze that the literary texts themselves often resist.

This paper proposes a corrective through what we term Sovereign AI Datasets (SADs): culturally anchored training corpora built from within Indian literary scholarship, guided by postcolonial theory, and designed to encode the interpretive principles that IEL itself demands. The term 'sovereign' is deliberate: it invokes the political philosophy of data sovereignty while simultaneously asserting the epistemological autonomy of non-Western literary traditions against their computational marginalisation.

The remainder of this paper is organised as follows. Section 2 reviews existing literature at the interface of NLP and literary studies, with particular attention to documented biases. Section 3 develops the theoretical framework—the Computational-Postcolonial Framework (CPF)—that grounds the SAD architecture. Section 4 presents the methodology for constructing SADs, including corpus design, annotation ontology, and validation protocol. Section 5 reports experimental results comparing SAD-trained models against Western-centric baselines. Section 6 discusses implications for computational humanities and future research directions. Section 7 concludes.

2. LITERATURE REVIEW

2.1 NLP and Literary Analysis: Capabilities and Constraints

Computational approaches to literary analysis have proliferated across a range of tasks: authorship attribution (Stamatatos, 2009), sentiment analysis (Liu, 2012), narrative structure modelling (Elsner, 2012), character network extraction (Elson et al., 2010), and topic modelling (Blei et al., 2003). These methods offer genuine analytical leverage, particularly for large corpora where close reading is impractical. However, as Underwood (2019) observes in Distant Horizons, computational models in literary studies are interpretive instruments, not transparent windows onto texts. Their outputs encode the assumptions of their training data as surely as any human critical tradition.

The problem of training data bias in NLP has been extensively documented in computational linguistics. Zhao et al. (2018) demonstrate systematic gender bias in word embeddings; Blodgett et al. (2020) offer a comprehensive survey of racial and dialectal bias in NLP systems, noting that models consistently underperform on African American Vernacular English. Joshi et al. (2020) map the severe imbalance in language resource availability, revealing that the top five languages—all European—account for over 90% of available NLP training data. Indian English, despite its estimated 125 million speakers, occupies a peripheral position within this hierarchy.

2.2 Indian English in Computational Linguistics

Research specifically addressing the computational treatment of Indian English is sparse and largely instrumental. Existing work focuses on automatic speech recognition calibrated for Indian English phonological patterns (Behravan et al., 2014), grammatical error correction for learner English (Rozovskaya & Roth, 2016), and code-switching detection in social media (Barman et al., 2014). These studies treat Indian English as a learner variety or a sociolinguistic curiosity rather than as a mature literary language with its own aesthetics.

Literary-computational work specifically targeting IEL is virtually absent from the literature. Studies that claim to analyse 'postcolonial literature' computationally—such as Wilkens (2012) on geographical imagination in American fiction, or Jockers' (2013) macro-analysis of nineteenth-century novels—remain anchored to Western canonical traditions. The postcolonial label is applied to analytical frame, not training corpus or annotation schema.

2.3 Postcolonial Theory and the Critique of Computational Epistemology

A growing body of critical scholarship within Digital Humanities has begun to interrogate the colonial assumptions embedded within digital methods. Noble (2018) in Algorithms of Oppression documents how search algorithms reproduce racial hierarchies; Benjamin (2019) in Race After Technology theorises 'the New Jim Code'—algorithmic systems that encode racial bias under the appearance of objectivity. In the context of world literature and Global South knowledge systems, Risam (2018) argues in New Digital Worlds that DH methods are shaped by postcolonial power asymmetries that determine whose texts, languages, and interpretive conventions are computationally legible.

Within Indian literary scholarship, Mukherjee (2000) traces the contested cultural politics of IEL, while Mehrotra (1992) documents the diverse regional and linguistic substrates that inflect Indian writing in English. Spivak's (1988) foundational question—'Can the subaltern speak?'—acquires a new computational dimension: can existing NLP systems hear subaltern textuality when the very frequencies they are trained to detect are calibrated to dominant literary traditions? We argue that they cannot, and that dataset sovereignty is the prerequisite for computational audibility.

2.4 Gaps in Existing Research

Three substantive gaps emerge from this review. First, no existing NLP framework has been systematically designed to process the specific linguistic and literary features of Indian English literature. Second, postcolonial theory has been applied as a critical meta-commentary on DH practice but rarely operationalised within the technical architecture of NLP systems. Third, the concept of data sovereignty in AI has been discussed in relation to national data governance and Indigenous data rights (Carroll et al., 2020) but not yet developed as a literary-computational methodology. This paper addresses all three gaps simultaneously.

3. THEORETICAL FRAMEWORK: THE COMPUTATIONAL-POSTCOLONIAL FRAMEWORK (CPF)

3.1 Foundational Theoretical Pillars

The Computational-Postcolonial Framework (CPF) synthesises three postcolonial theoretical traditions with computational methodology, producing a hybrid analytical architecture that is simultaneously critical and operational.

The first pillar draws on Homi Bhabha's (1994) concept of cultural hybridity and the third space of enunciation. In The Location of Culture, Bhabha argues that colonial discourse is never simply an imposition but is always already contaminated by the ambivalence of the colonised response. Indian English literature inhabits precisely this third space: it deploys the language of colonial administration while simultaneously subverting its semantic authority. For NLP, this means that lexical forms must not be interpreted through their denotative function in standard British English but must be read relationally, against their postcolonial semantic displacement.

The second pillar is drawn from Gayatri Chakravorty Spivak's (1988) theorisation of the subaltern. Spivak's argument that the subaltern cannot speak within hegemonic discursive structures applies directly to the computational treatment of marginalised literary voices. Dalit literature, women's narratives of partition, and working-class fiction written in IEL have been systematically excluded from canonical corpora; they are computationally subaltern not because they do not exist but because the dominant training datasets do not register their existence. The CPF incorporates subaltern inclusion as an architectural principle, not an optional supplement.

The third pillar derives from Frantz Fanon's (1961, 1967) analysis of the psychopathological dimensions of colonial language. Fanon's insight that the colonised writer experiences the colonial language as simultaneously the medium of expression and the instrument of psychic alienation illuminates the affective complexity of IEL. A text's use of English may simultaneously perform compliance, conceal resistance, and enact self-fashioning. Standard sentiment analysis, which assigns binary or scalar valences to linguistic features, is structurally unable to process this affective multiplicity. The CPF requires multi-valent affective annotation that can capture Fanonian ambivalence.

3.2 Operationalising the CPF

The CPF translates these theoretical commitments into five operational principles for SAD construction:
Principle 1 — Epistemic Pluralism: Training corpora must include texts produced within Indian literary traditions across regional, caste, gender, and religious axes, ensuring that no single literary-cultural formation is taken as the normative template.
Principle 2 — Relational Annotation: Annotation schemas must encode meaning relationally rather than absolutely. A word's semantic value is determined not by its dictionary definition in British English but by its function within the specific postcolonial discourse of the text.
Principle 3 — Multilingual Substrate Sensitivity: IEL annotation must account for the deep structural influence of Indian language substrates on English-medium texts. Syntactic patterns that deviate from Standard English norms are not errors but features of a distinct literary register.
Principle 4 — Affective Complexity: Sentiment and affect annotation must move beyond binary positive/negative or scalar schemes to accommodate irony, ambivalence, code-switched emotional registers, and subaltern affect.
Principle 5 — Interpretive Accountability: All annotation decisions must be traceable to explicit theoretical and cultural rationales, ensuring that the dataset's interpretive assumptions are transparent and open to scholarly contestation.

4. METHODOLOGY

4.1 Corpus Architecture for Sovereign AI Datasets

The SAD corpus is structured around four thematic-historical strata that correspond to the major phases of IEL's development: (i) early colonial and anti-colonial writing (1880–1947), (ii) postindependence national literature (1947–1975), (iii) the generation of global Indian fiction (1975–2000), and (iv) twenty-first century diasporic, Dalit, and feminist writing (2000–present). Each stratum is designed to represent the dominant literary concerns, formal experiments, and ideological tensions of its period.

Text selection within each stratum is guided by three criteria. The first is canonical diversity: texts are selected to represent the full range of regional, linguistic, and social positions within IEL, including works that have been historically marginalised by Western publishing markets. The second is formal heterogeneity: the corpus includes novels, short stories, poetry, autobiography, and essay, acknowledging that computational models trained exclusively on one form impose formal assumptions on others. The third is linguistic density: priority is given to texts that exhibit the highest concentration of the linguistic features most resistant to standard NLP processing—code-switching, dialectal variation, mythological reference, and postcolonial irony.

Table 1: SAD Corpus Composition by Stratum and Feature Type

Stratum	Period	No. Texts	Tokens (approx.)	Key Literary Features
Early Colonial/Anti-colonial	1880–1947	42	1.2M	Colonial mimicry, nationalist allegory, Orientalist resistance
Post-independence National	1947–1975	67	2.8M	Nation-building discourse, partition trauma, caste critique
Global Indian Fiction	1975–2000	89	4.5M	Diasporic hybridity, magic realism, market-oriented cosmopolitanism
Contemporary Polyphonic	2000–present	103	5.1M	Dalit autobiography, feminist counter-narrative, digital multilingualism

4.2 Annotation Ontology

The SAD annotation ontology is organised across six dimensions, each addressing a distinct feature of IEL that standard NLP annotation fails to capture adequately.

Dimension 1 — Linguistic Register: Texts are annotated for diglossic register at the sentence level, distinguishing among high-formal colonial English, middle-register educated Indian English, vernacular-inflected speech, and code-switched utterances. This enables register-sensitive processing rather than the false normalisation imposed by standard tokenisers.

Dimension 2 — Cultural Entity Classification: Named entity recognition (NER) labels are extended to include Indian cultural entity types absent from standard ontologies: mythological-scriptural entities (MSE), caste-community markers (CCM), vernacular toponym variants (VTV), and festival-ritual references (FRR).

Dimension 3 — Intertextual Mapping: Passages exhibiting intertextual resonance with Indian oral, classical, or folk traditions are annotated with source-tradition tags, enabling models to process allusion as a semantic device rather than miscategorising it as proper noun usage or metaphor.

Dimension 4 — Postcolonial Affect: A novel affect taxonomy replaces the standard Ekman-derived emotion labels with categories derived from postcolonial affective theory: compliant resistance, subaltern grief, diasporic longing, ironic solidarity, and anti-colonial rage.

Dimension 5 — Syntactic Substrate Marking: Sentences exhibiting substrate influence from specific Indian languages (Hindi-Urdu, Tamil, Bengali, Malayalam, Marathi) are annotated accordingly, allowing models to distinguish postcolonial syntactic variation from grammatical error.

Dimension 6 — Narrative Agency: Character action annotations are extended to include subaltern agency markers that distinguish between surface compliance and underlying resistance—a distinction invisible to standard event-detection models trained on Western-canonical fiction.

Table 2: Comparison of Standard NLP Ontology vs. SAD Annotation Ontology

Annotation Dimension	Standard NLP Label	SAD Label	Theoretical Basis
Entity Classification	PERSON, ORG, LOC	MSE, CCM, VTV, FRR	Cultural ontology (Bhabha)
Sentiment	Positive / Negative / Neutral	6-category postcolonial affect	Fanon; Spivak
Language Variety	Error / Non-standard	Substrate-marked register	Kachru's Concentric Circles
Narrative Function	Agent / Patient	Agency-spectrum (5 levels)	Subaltern theory
Intertextuality	Metaphor / Allusion	Tradition-sourced intertextual tag	Menon; Dharwadker
Code-switching	Foreign word	Diglossic register marker	Mukherjee; Ashcroft et al.

4.3 Annotation Validation Protocol

Annotation quality is ensured through a three-stage validation protocol. In stage one, all annotators—drawn from postgraduate students in Indian literary studies at universities across five regional zones—complete a training programme that integrates postcolonial theory with annotation practice. In stage two, a stratified random sample of 15% of all annotated passages is reviewed by an expert panel comprising literary scholars specialising in the relevant regional tradition, computational linguists with NLP annotation experience, and a cultural consultant for each substrate language. Inter-annotator agreement is measured using Cohen's Kappa for categorical labels, with a minimum acceptance threshold of $\kappa \geq 0.75$.

Stage three involves a dialectical revision procedure: passages where annotation disagreement exceeds threshold are returned to the original annotators with the expert panel's commentary, and the resulting discussion is archived as part of the dataset's interpretive provenance record. This procedure ensures that annotation is not mechanically imposed but theoretically grounded, and that the dataset retains a scholarly audit trail consistent with the principles of interpretive accountability articulated in the CPF.

4.4 Model Training Configuration

Prototype SAD-trained models are developed using a fine-tuning approach applied to a multilingual pre-trained transformer architecture (mBERT, Devlin et al., 2019) as the base model, fine-tuned on the SAD corpus with task-specific classification heads for each of the six annotation dimensions. Training is conducted on a held-out 80/10/10 split (train/validation/test). Hyperparameter optimisation is performed using Bayesian search over learning rate (1e-5 to 5e-4), batch size (16–64), and dropout rate (0.1–0.4). Three baseline models are established for comparative evaluation: (i) a vanilla mBERT model with no fine-tuning, (ii) mBERT fine-tuned on the BooksCorpus and Wikipedia corpus, and (iii) a commercially available literary NLP system trained on Western-canonical fiction.

5. RESULTS AND ANALYSIS

5.1 Quantitative Performance Metrics

Table 3 presents comparative performance metrics across four tasks for SAD-trained models versus all three baseline conditions. Results are reported as F1-scores (macro-averaged) on the test partition. All differences between SAD and baselines are statistically significant at $p < 0.01$ (paired t-test, n=50 text samples per stratum).

Table 3: F1-Score Comparison Across Tasks (SAD vs. Baselines)

Task	Baseline-1 (Vanilla mBERT)	Baseline-2 (BookCorpus mBERT)	Baseline-3 (Commercial)	SAD Model	Improvement vs. Best Baseline
Sentiment/Affect Classification	0.521	0.584	0.601	0.714	+18.7%
Narrative Role Classification	0.488	0.539	0.563	0.689	+22.3%
Cultural Metaphor Recognition	0.312	0.371	0.388	0.510	+31.4%
Substrate Syntax Identification	0.274	0.319	0.344	0.617	+79.4%
Intertextual Reference Detection	0.401	0.448	0.471	0.629	+33.5%

The most pronounced improvements are observed in substrate syntax identification (+79.4%) and cultural metaphor recognition (+31.4%)—precisely the tasks that require the deepest cultural-linguistic grounding. These results validate the hypothesis that baseline models' underperformance in IEL analysis is not a general limitation of transformer architectures but a specific consequence of training data bias: when models are given culturally appropriate training signal, their performance improves dramatically.

5.2 Qualitative Case Analysis

5.2.1 Postcolonial Irony in Anand's Untouchable

Mulk Raj Anand's Untouchable (1935) presents a paradigmatic case of the annotation failures characteristic of Western-centric models. The sentence 'Bakha thanked the sahib and felt warm inside with the compliment, even as his back ached from the blow' is classified by Baseline-2 as expressing 'positive sentiment (gratitude).' The SAD model, trained on annotations informed by Fanonian affect theory, correctly classifies it as 'subaltern deference masking traumatic subjection'—a reading that the text's narrative context unambiguously supports. The discrepancy illustrates how standard sentiment analysis, calibrated to sincerity conventions of Western psychological realism, systematically misreads the performative double consciousness that characterises colonial-era IEL.

5.2.2 Mythological Intertextuality in Roy's The God of Small Things

In Arundhati Roy's The God of Small Things (1997), the figure of Pappachi's moth is consistently misclassified by all baseline models as a 'zoological reference' or 'minor motif.' The SAD model, informed by annotations drawn from Malayalam literary scholarship, correctly identifies it as an intertextual condensation of the Narakasura myth and its associated cultural complex of shame, ancestral sin, and patriarchal haunting. This classification enables downstream analysis to correctly situate the moth within the novel's overarching meditation on caste, gender, and the 'Love Laws'—a reading that the baseline models' entity classification schemas structurally preclude.

5.2.3 Dalit Autobiography and Subaltern Agency

Baby Kamble's The Prisons We Broke (originally written in Marathi, translated into English by Maya Pandit, 2008) presents a particularly illuminating test case for narrative agency classification. Baseline models consistently classify the protagonist's acts of ritual compliance—performing caste-prescribed labour, observing untouchability restrictions—as 'patient' roles in the narrative schema, thereby reproducing the caste ideology the autobiography is written to critique. The SAD model's five-level agency spectrum correctly identifies these acts as 'constrained agency within systemic subjection,' a classification that preserves the text's political argument and enables computational analysis to contribute to rather than contradict postcolonial literary interpretation.

6. DISCUSSION

6.1 Dataset Sovereignty as Epistemic Decolonisation

The results presented in Section 5 demonstrate that the SAD approach produces measurable computational improvements across all tested tasks. However, this paper argues that the significance of dataset sovereignty cannot be reduced to performance metrics. The deeper argument is epistemological: the construction of SADs enacts a form of computational decolonisation by refusing the tacit assumption that Anglo-American literary conventions constitute a universal interpretive baseline.

This argument has precedents in cognate fields. Carroll et al.'s (2020) CARE Principles for Indigenous Data Governance—Collective Benefit, Authority to Control, Responsibility, Ethics—offer a parallel framework in the context of Indigenous data sovereignty. The SAD project extends this logic to the literary-computational domain: Indian English literature has the collective right to benefit from computational analysis, the epistemic authority to determine the frameworks through which it is analysed, the right to demand responsible annotation practices, and the right to expect that AI systems engaging with it do so ethically rather than reductively.

6.2 Limitations and Ethical Considerations

Several limitations of the present study require acknowledgement. First, the prototype SAD corpus, while diverse across regional and historical axes, cannot claim comprehensive coverage of IEL's full range. Oral literary traditions, performance texts, and informal digital writing in Indian English remain underrepresented. Second, the annotation ontology, despite its theoretical sophistication, involves interpretive choices that are themselves culturally situated: the scholars who developed the ontology bring their own positionalities and blind spots to the annotation task. Third, the use of mBERT as the base architecture imposes architectural assumptions derived from multilingual but Western-dominated pre-training that the SAD fine-tuning only partially overcomes.

On the ethical dimension, the construction of literary AI training datasets raises issues of intellectual property, textual consent, and the commodification of cultural heritage. This project is committed to open-access distribution of the SAD corpus for non-commercial scholarly use, with attribution protocols that acknowledge the literary scholars whose interpretive labour underlies the annotation schema.

6.3 Implications for Computational Humanities

The CPF and the SAD methodology have implications that extend beyond Indian English literature. Any computational approach to literary traditions that have been historically marginalised within world literature scholarship—African, Caribbean, Southeast Asian, Latin American—faces structurally analogous problems: the imposition of Western narrative templates, the invisibility of culturally specific literary devices, and the reproduction of colonial hierarchies of literary value through algorithmic proxies. The SAD methodology offers a transferable framework for building culturally sovereign datasets in any literary tradition, provided that equivalent investments are made in theoretically grounded, community-accountable annotation practices.

The paper also contributes to debates within the emerging field of Critical AI Studies. Floridi et al.'s (2018) 'AI4People' framework calls for AI development that is beneficial, fair, transparent, and accountable. The CPF instantiates all four principles within the literary domain: SADs are more beneficial for IEL analysis, fairer in their representation of marginalised literary traditions, transparent in their annotation rationale, and accountable through their provenance records.

7. CONCLUSION

This paper has argued that the computational analysis of Indian English literature requires not merely better algorithms but better datasets—corpora built from within the cultural and theoretical frameworks that IEL itself inhabits. The Sovereign AI Dataset model, grounded in the Computational-Postcolonial Framework synthesised from Bhabha, Spivak, and Fanon, offers a systematic methodology for constructing such corpora. Experimental results demonstrate that SAD-trained models outperform Western-centric baselines across all tested literary analysis tasks, with the largest gains precisely in those tasks—substrate syntax identification, cultural metaphor recognition, postcolonial affect classification—where cultural grounding matters most.

The concept of dataset sovereignty reframes the politics of AI in the humanities: it insists that the authority to determine how a literary tradition is computationally read belongs not to the global technology industry but to the scholarly communities that have produced, debated, and interpreted that tradition. This is not an argument against computational methods in literary studies; it is an argument for computational methods that are epistemically accountable to the literatures they claim to analyse.

Future work will focus on expanding the SAD corpus to full-scale coverage of IEL across all four strata, developing SAD methodologies for allied South Asian literary traditions in translation, and integrating the CPF into graduate curricula in Digital Humanities in Indian universities. The ultimate aim is a computational literary studies that can read the world's literature in the world's own terms.

REFERENCES

1. Anand, M. R. (1935). Untouchable. Wishart Books.
2. Ashcroft, B., Griffiths, G., & Tiffin, H. (2002). The Empire Writes Back: Theory and Practice in Post-Colonial Literatures (2nd ed.). Routledge.
3. Barman, U., Das, A., Wagner, J., & Foster, J. (2014). Code mixing: A challenge for language identification in the language of social media. In Proceedings of the First Workshop on

Computational Approaches to Code Switching (pp. 13–23). Association for Computational Linguistics.
4. Behravan, H., Hautamäki, V., & Kinnunen, T. (2014). Foreign accent detection from spoken English using i-vectors and Gaussian backend. INTERSPEECH 2014, 2396–2400.
5. Benjamin, R. (2019). Race After Technology: Abolitionist Tools for the New Jim Code. Polity Press.
6. Bhabha, H. K. (1994). The Location of Culture. Routledge.
7. Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3, 993–1022.
8. Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of 'bias' in NLP. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 5454–5476).
9. Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., ... & Anderson, J. (2020). The CARE Principles for Indigenous Data Governance. Data Science Journal, 19(1), 43.
10. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of NAACL-HLT 2019 (pp. 4171–4186).
11. Elsner, M. (2012). Character-based kernels for novelistic plot structure. In Proceedings of EACL 2012 (pp. 634–644).
12. Elson, D. K., Dames, N., & McKeown, K. R. (2010). Extracting social networks from literary fiction. In Proceedings of ACL 2010 (pp. 138–147).
13. Fanon, F. (1961). The Wretched of the Earth (C. Farrington, Trans.). Grove Press.
14. Fanon, F. (1967). Black Skin, White Masks (C. L. Markmann, Trans.). Grove Press.
15. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). AI4People—An ethical framework for a good AI society. Minds and Machines, 28(4), 689–707.
16. Jockers, M. L. (2013). Macroanalysis: Digital Methods and Literary History. University of Illinois Press.
17. Joshi, P., Santy, S., Budhiraja, A., Bali, K., & Choudhury, M. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. In Proceedings of ACL 2020 (pp. 6282–6293).
18. Kamble, B. (2008). The Prisons We Broke (M. Pandit, Trans.). Orient BlackSwan. (Original work published 1986)
19. Kachru, B. B. (1986). The Alchemy of English: The Spread, Functions and Models of Non-native Englishes. Pergamon Press.
20. Liu, B. (2012). Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers.
21. Mehrotra, A. K. (Ed.). (1992). A History of Indian Literature in English. Columbia University Press.
22. Mukherjee, M. (2000). The Perishable Empire: Essays on Indian Writing in English. Oxford University Press.
23. Noble, S. U. (2018). Algorithms of Oppression: How Search Engines Reinforce Racism. New York University Press.
24. Risam, R. (2018). New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy. Northwestern University Press.
25. Roy, A. (1997). The God of Small Things. HarperCollins.
26. Rozovskaya, A., & Roth, D. (2016). Grammatical error correction: Machine translation and classifiers. In Proceedings of ACL 2016 (pp. 2205–2215).
27. Spivak, G. C. (1988). Can the subaltern speak? In C. Nelson & L. Grossberg (Eds.), Marxism and the Interpretation of Culture (pp. 271–313). University of Illinois Press.
28. Stamatatos, E. (2009). A survey of modern authorship attribution methods. Journal of the American Society for Information Science and Technology, 60(3), 538–556.
29. Underwood, T. (2019). Distant Horizons: Digital Evidence and Literary Change. University of Chicago Press.
30. Wilkens, M. (2012). The geographic imagination of Civil War-era American fiction. Literary and Linguistic Computing, 28(2), 238–253.

31. Zhao, J., Wang, T., Yatskar, M., Ordonez, V., & Chang, K. W. (2018). Gender bias in coreference resolution: Evaluation and debiasing methods. In Proceedings of NAACL-HLT 2018 (pp. 15–20).