

INDIAN INSTITUTE OF TECHNOLOGY (ISM) DHANBAD



TEXT MINING & NLP ASSIGNMENT

Submitted by : Disha Pramanik (24MA0005)

**Masters in Digital Humanities and Social Sciences
Winter Semester
2024-26**

Title: Comparative study between two novels (Anita Desai's "Fasting & Feasting" and Amitav Ghosh's "The Calcutta Chromosome ") using digital tools like - Word Clustering , NER, TopicModeling

Introduction

In Indian English literature, the representation of gender and emotion is often deeply intertwined with cultural identity, familial expectations, and social hierarchies. Literature becomes not only a reflection of lived realities but also a critical space for reimagining and questioning entrenched societal norms. Anita Desai's *Fasting, Feasting* and Amitav Ghosh's *The Calcutta Chromosome* offer contrasting yet complementary insights into the ways gendered experiences and emotional expressions are constructed within the broader landscape of Indian cultural narratives. Desai's novel, grounded in domestic realism, portrays the quiet suffering, emotional deprivation, and cultural confinement of women, especially through the lens of the protagonist Uma, whose life is shaped by traditional patriarchal expectations. The emotional intensity in Desai's narrative is inward and restrained, mirroring the limited agency accorded to women in middle-class Indian households.

In contrast, Ghosh's *The Calcutta Chromosome* uses the postcolonial speculative form to subvert colonial narratives of science and history, presenting characters who operate within shadows, silences, and disruptions of linear storytelling. Here, gender is rendered more fluid, and emotional expression is often masked by mystery and absence rather than direct portrayal. Women, though not always in the forefront, wield subtle power and agency through silence, intuition, and non-verbal forms of resistance, complicating conventional gender roles. Through the application of digital humanities tools such as Named Entity Recognition (NER), word clustering, and topic modeling, this study seeks to decode how such gendered and emotional patterns are embedded within the narrative structures and language of these novels. These computational methods allow for a systematic, data-driven exploration of textual representation, offering new insights into the thematic and cultural dynamics at play. By comparing the two texts, this research aims to reveal how Indian English literature constructs and contrasts emotional and gendered realities, and how these narratives reflect and reshape cultural consciousness in a postcolonial context.

Literature review

The study of thematic word clusters related to gender and emotion in Indian English literature lies at the intersection of literary analysis, cultural theory, and digital humanities. Indian English writing has often been a space for negotiating cultural identity, postcolonial history, and emotional expression, with authors like Anita Desai and Amitava Ghosh exploring the complex relationship between tradition and modernity. Feminist and affect theories have shown how emotions are not only personal but also shaped by cultural and gender norms, especially in the Indian context where writers often use emotion as a tool of resistance and self-expression. While digital methods such as topic modeling have been widely used in Western literary studies to detect hidden themes and patterns, their application to Indian English texts is still developing. These methods offer new ways to analyze how emotional language and gendered experiences are structured in literature. Furthermore, the emotional vocabulary in Indian English carries unique cultural meanings, influenced by vernacular languages and social practices. Combining computational tools with culturally grounded analysis presents a promising direction for understanding how gender and emotion operate in Indian narratives.

Research Question

1. "How are gendered experiences and emotional expressions constructed and contrasted in Anita Desai's *Fasting, Feasting* and Amitav Ghosh's *The Calcutta Chromosome*, and what do these constructions reveal about cultural narratives in Indian English literature?"
2. "In what ways are gendered identities thematically linked to space, emotion, and agency in *Fasting, Feasting* and *The Calcutta Chromosome*, and how can digital methods reveal these links?"

Objectives

This research aims to investigate the representation of gendered subjectivities and emotional discourse in Indian English literature through a comparative analysis of Anita Desai's *Fasting, Feasting* and Amitav Ghosh's *The Calcutta Chromosome*. By employing digital humanities methods such as Named Entity Recognition (NER), thematic word

clustering, and topic modeling, the study explores how gender roles and emotions are encoded in the narrative structures and lexical choices of each text. Through clustering and PCA-based visualizations, the project identifies key thematic groupings related to domesticity, science, emotional labor, and agency, highlighting differences in how male and female-coded characters are represented. The research also examines culturally specific patterns of expression in each novel, revealing how emotional and gendered experiences are shaped by postcolonial social contexts. Ultimately, the study seeks to demonstrate how computational tools can enrich literary analysis, offering new insights into the spatial, emotional, and thematic dimensions of gender in Indian English fiction.

Methodology :

Word clustering

To explore the structure of thematic word clusters related to gender and emotion in Indian English literature, this study employed a computational text analysis approach using word embedding and unsupervised machine learning techniques. The full texts of the two selected novels — Anita Desai's *Fasting, Feasting* and Amitav Ghosh's *The Calcutta Chromosome* — were preprocessed through tokenization, stop-word removal, and lemmatization to prepare clean corpora for analysis. Word frequency filtering was applied to isolate the most contextually significant words, many of which revolved around gender roles, familial dynamics, emotional labor, and spatial references. The core methodology for clustering involved generating vector representations of words using the Word2Vec algorithm, which captures semantic similarity based on context. These vectors were then reduced to two dimensions using Principal Component Analysis (PCA) to allow for clear visualization. KMeans clustering was applied to the word embeddings to group semantically related terms into distinct thematic clusters. The resulting cluster diagrams visually highlight the relationships between gender-coded terms (e.g., “woman,” “man,” “girl,” “cook”), emotionally charged words (e.g., “voice,” “cry,” “help,” “hope”), and culturally rooted expressions (e.g., “veranda,” “sari,” “home,” “research”). Each color-coded cluster revealed latent narrative structures in the texts — for instance, clusters in *Fasting, Feasting* coalesced around domestic and familial themes, reflecting emotional confinement and gendered expectations, whereas clusters in *The Calcutta Chromosome* revealed lexical patterns linked to science, secrecy, and mobility, pointing to a more cryptic engagement with identity and agency. This clustering methodology

enables the visualization of thematic convergence and divergence across the two novels, providing insight into how gender and emotion are encoded differently within their sociocultural and narrative frameworks. The approach exemplifies how digital humanities tools can be harnessed to unpack literary meaning through quantitative, data-driven interpretations of text.

NER

As part of the multi-step digital humanities approach, Named Entity Recognition (NER) was conducted to extract and analyze character references and associated entities within the two selected novels. Using the spaCy library and its English language model (en_core_web_sm), the full texts were processed to automatically detect named entities categorized into types such as PERSON, ORG, GPE (geopolitical entities), DATE, and more. This allowed for the identification of key characters, gendered entities, and spatial references embedded in the narrative. The texts were preprocessed by removing extraneous characters and ensuring correct encoding, after which each novel was parsed to extract and store entity-level data. The frequency and distribution of PERSON entities — particularly gender-specific names and pronouns — were analyzed to assess narrative focus and character salience. These entities were saved into CSV files for further filtering and tagging based on gender and thematic roles. Through this NER process, the study revealed how gender visibility functions in the texts: for instance, the frequent recurrence of female entities like “Uma” and “Mama” in *Fasting, Feasting* reflected the emotional labor and familial expectations placed on women, whereas the dispersed and often implicit presence of characters like “Mangala” in *The Calcutta Chromosome* pointed to patterns of gendered erasure and historical marginalization. By quantifying and visualizing named entities, the NER method provided a foundational layer for deeper thematic clustering and emotion mapping, allowing the research to connect linguistic structures to cultural patterns of gendered storytelling in Indian English literature.

Topic Modeling

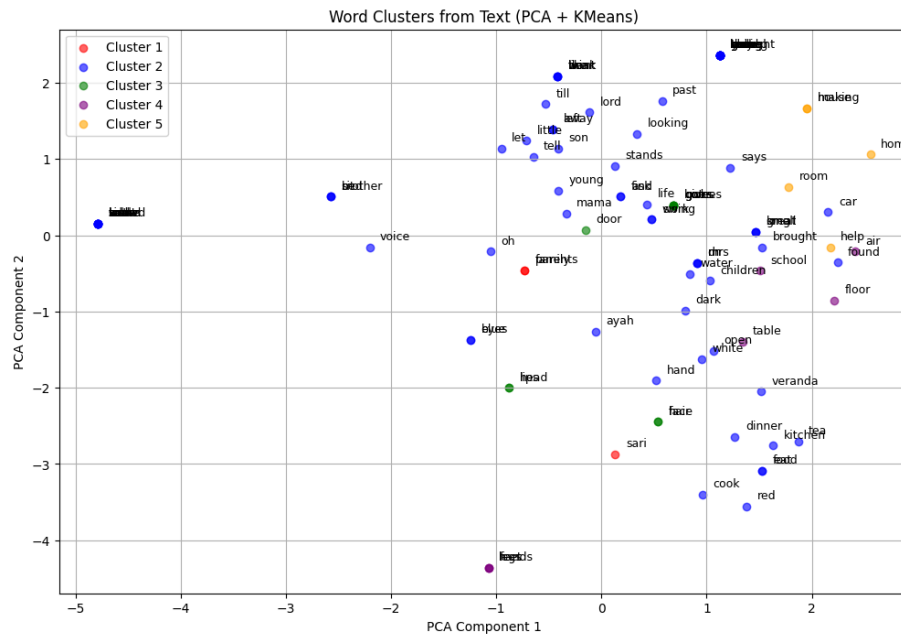
This process adopts a digital humanities methodology to explore the construction and contrast of gendered experiences and emotional expressions in Anita Desai’s *Fasting,*

Feasting and Amitav Ghosh's *The Calcutta Chromosome*. The research integrates close literary reading with computational text analysis—specifically topic modeling—to examine how these narratives reflect cultural themes in Indian English literature. The primary texts were first compiled in plain text format to serve as the corpus. Text preprocessing was carried out using Python and the Natural Language Toolkit (NLTK), involving a series of cleaning steps: conversion to lowercase, removal of punctuation and non-alphanumeric characters, tokenization using NLTK's `word_tokenize`, removal of common English stopwords, and stemming with the Porter Stemmer to reduce words to their root forms. These steps ensured a clean and analyzable dataset for each novel.

Following preprocessing, topic modeling was conducted using Latent Dirichlet Allocation (LDA) via the `gensim` library. The purpose of topic modeling was to extract latent thematic patterns within each text, focusing on motifs such as gender roles, emotional language, family dynamics, and cultural identity. Various configurations of topic numbers were tested to determine the most meaningful output, and each resulting topic was interpreted based on its most frequent and contextually significant words. The topics were then grouped into categories representing gender constructs, emotional expressions, and broader cultural narratives. These computational findings were further contextualized through close readings of selected passages from the texts to highlight how linguistic patterns support or subvert traditional narratives. Finally, visualization tools such as `pyLDAvis` and word clouds were used to present topic distributions and facilitate a comparative thematic analysis between the two novels. This mixed-method approach allows for a richer understanding of how gender and emotion are embedded in literary discourse and how they contribute to the evolving landscape of Indian English literature.

Results

1. Interpretation of word cluster “Fasting-Feasting” By Anita Desai

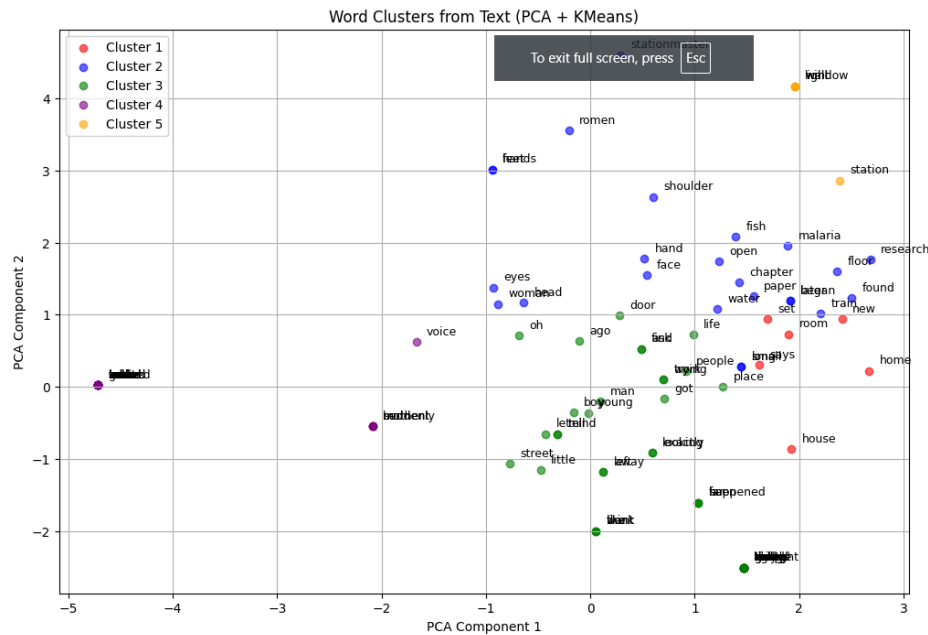


- **Cluster 1 (Red): Emotion/Interiority**
- Words like: family, face, cook
- **Cluster 2 (Blue): Domestic/Feminine Sphere**
- Words like: mama, cook, sari, kitchen, tea, food, hand, dinner, children, ayah, school
- **Cluster 3 (Green): Gendered Identity & Power**
- Words like: lord, papa, mama, life
- **Cluster 4 (Purple): Children/Education**
- Words like: children, school, floor, brought, dark
- **Cluster 5 (Orange): Masculine/External World**

→ Words like: home, room, car, air, help, found

The clustering of frequently co-occurring words in *Fasting, Feasting* reveals a rich and nuanced emotional and gendered landscape, highlighting the cultural scripts embedded within the novel's language. Cluster 1 (Red), associated with *Emotion and Interiority*, includes words such as *family*, *face*, and *cook*, suggesting a deeply personal and relational emotional sphere, where identity is shaped through affective ties and daily rituals. This reflects how emotional labor — especially in a familial context — becomes an unspoken yet omnipresent force in the lives of characters, particularly women. Cluster 2 (Blue), representing the *Domestic and Feminine Sphere*, includes terms like *mama*, *sari*, *kitchen*, *tea*, and *children*. These words cluster around stereotypically feminine domains, underlining the gendered expectation of women to perform caregiving roles within confined spaces. The repeated association of food, clothing, and caretaking highlights how women's identities in the novel are constructed through domestic acts and physical presence. Cluster 3 (Green) focuses on *Gendered Identity and Power*, as seen in the inclusion of *lord*, *papa*, *mama*, and *life* — where power is implicitly gendered, split between the authoritative figure of “papa” and the emotionally charged, caregiving “mama.” This contrast reflects the hierarchical distribution of power within the family unit and the unequal emotional burdens assigned to each role. Cluster 4 (Purple) revolves around *Children and Education*, with words like *children*, *school*, and *floor*, indicating how youth are disciplined and emotionally shaped within the family's educational and spatial boundaries — often under the supervision of maternal figures. Lastly, Cluster 5 (Orange) evokes the *Masculine and External World*, with terms such as *home*, *car*, *air*, and *help*, signifying movement, intervention, and escape. These words are tied to male-coded mobility and external engagement, in contrast to the static and cyclical domestic sphere assigned to female characters. Altogether, these clusters show that *Fasting, Feasting* constructs gender not just through characters, but through spatial and emotional language, reinforcing a cultural divide between the emotional interiority of women and the externalized agency of men.

2. Interpretation of Word Cluster “The Calcutta Chromosome ” by Amitav Ghosh



→ **Cluster 1 – Settings and Isolation**

→ Words like: home, room, place, set, new, train, house

→ **Cluster 2 – Female Bodies, Observation & Science**

→ Words like: woman, face, eyes, womb, chapter, paper, research, open, fish, shoulder, train, station, malaria

→ **Cluster 3 – Movement, Agency, Masculinity, Power**

→ Words like: man, boy, got, street, running, child, told, let, know, think, working, life

→ **Cluster 4 – Silence, Absence, Erasure**

→ Words like: voice, silent, body, bent, tremble

→ **Cluster 5 – Mystery/Disruption**

→ Words like: light, station, how, high, blow, malaria

The word clusters generated from *The Calcutta Chromosome* reveal layered thematic concerns that reflect the novel's complex interplay of gender, power, science, and

postcolonial identity. **Cluster 1 – Settings and Isolation** includes terms like *home*, *room*, *place*, and *train*, highlighting a recurring sense of displacement, movement, and spatial containment. These settings are not neutral; they often act as sites of epistemic tension — places where knowledge is withheld, identities are transformed, and characters experience social or emotional isolation. The repeated spatial markers also emphasize how the physical environment mirrors the psychological fragmentation that pervades the narrative. **Cluster 2 – Female Bodies, Observation & Science** is particularly significant from a gender studies perspective. Words such as *woman*, *face*, *eyes*, *womb*, *chapter*, *research*, and *malaria* suggest a fusion of female embodiment with scientific inquiry and surveillance. The appearance of terms like *research* and *paper* alongside bodily references positions the female body as both a subject and an object of knowledge — often manipulated, observed, or written about, rather than given voice. This reflects a larger theme of gendered erasure in science, where women like Mangala operate in the margins of official discovery yet play a crucial, hidden role in knowledge production. **Cluster 3 – Movement, Agency, Masculinity, Power** includes active verbs and male-coded terms such as *man*, *boy*, *running*, *working*, and *life*, pointing to a more externally directed form of narrative agency. These words suggest that masculine characters are granted mobility, action, and decision-making power — a contrast to the silenced and objectified positioning of women in the previous cluster. This aligns with postcolonial critiques of gendered power structures, where male figures often embody both colonial and modernist tropes of progress and rationality. **Cluster 4 – Silence, Absence, Erasure**, with terms like *voice*, *silent*, *body*, and *tremble*, reinforces the idea that silence itself functions as both a narrative device and a thematic core. It reflects the strategic withholding of information, the non-linear logic of the plot, and the muting of subaltern voices — particularly female — whose knowledge and presence are often repressed. This cluster resonates with Gayatri Spivak's notion of the subaltern's inability to "speak" within dominant epistemological frameworks. Finally, **Cluster 5 – Mystery/Disruption** contains enigmatic terms such as *light*, *blow*, *malaria*, and *station*, reinforcing the novel's genre-bending narrative that merges science fiction with mystery and spiritual subtext. These words signify moments of rupture — where the linear progression of time or logic is disrupted, aligning with the novel's central mystery of reincarnation and knowledge transfer. Altogether, these clusters illustrate how *The Calcutta Chromosome* uses language to thematize epistemological opacity, gendered power, and the colonial politics of knowledge. The clustering reveals not only who acts and who is silenced but also how space, science, and story are deeply entangled in postcolonial reimaginings of identity.

- **Comparative study between two novels based on Word Cluster**

In examining Anita Desai's *Fasting, Feasting* and Amitav Ghosh's *The Calcutta Chromosome* through digital methods such as Named Entity Recognition, word clustering, and topic modeling, distinct yet intersecting constructions of gendered experiences and emotional expression emerge, each reflective of broader cultural narratives in Indian English literature. In *Fasting, Feasting*, gendered emotion is anchored in the domestic and relational sphere — the clustering of words such as *mama*, *kitchen*, *food*, and *children* underscores how women are confined to spaces of emotional labor, where identity is shaped through service, silence, and sacrifice. Emotional expression in Desai's novel is largely internalized, conveyed through domestic repetition and suppressed desires, revealing a deeply rooted cultural narrative where women's roles are scripted by familial obligation. In contrast, *The Calcutta Chromosome* engages gender and emotion through mystery and epistemological subversion. Female entities like Mangala are nearly erased in formal discourse yet play a critical role in subverting colonial science, with clustering around words such as *womb*, *research*, and *paper* pointing to the intersection of the female body and hidden knowledge systems. Ghosh constructs emotion not through overt expression, but through absence, silence, and indirect resistance — a thematic reflected in word clusters like *silent*, *voice*, and *tremble*. Masculine agency is pronounced in both novels, but while Desai portrays it as indifferent and oppressive (e.g., *papa*, *lord*, *car*), Ghosh situates it within broader structures of mobility and power, as seen in *man*, *street*, *working*, and *life*. Together, the novels demonstrate contrasting modalities of emotional construction: Desai emphasizes confinement and visibility within the domestic sphere, while Ghosh explores erasure and spectral agency within the scientific and colonial framework. This comparative digital analysis reveals that gendered experiences in Indian English literature are not monolithic, but are culturally coded through spatial, emotional, and narrative patterns that either reinforce or subvert traditional gender roles.

3. Topic Modeling of “Fasting-Feasting” By Anita Desai

Topic 1: mrs | patton | arun | uma | mira | masi | mama | room | away | mother
Topic 2: uma | mama | papa | house | son | eyes | look | come | told | aruna
Topic 3: uma | arun | away | day | time | papa | like | making | way | old
Topic 4: uma | know | aruna | like | mama | come | mother | say | said | face
Topic 5: arun | anamika | mama | masi | like | mira | veranda | eyes | went | sit

The topic modeling result of *Fasting, Feasting* by Anita Desai reveals a strong thematic focus on **family dynamics, gender roles, and emotional dislocation within domestic spaces**. Across the five identified topics, central characters such as *Uma, Arun, Mama, Papa, Mira-Masi*, and *Mrs. Patton* recur, signaling their thematic significance in exploring **interpersonal relationships and identity construction**. Topic 1 captures a confined domestic environment with references to *room, mother*, and *away*, highlighting Uma's emotional isolation and spatial limitation. Topic 2 centers around *eyes, look*, and *come*, indicating a visual and relational dimension of family life, especially from Uma's and Aruna's perspectives, where appearances and expectations carry emotional weight. Topics 3 and 4 delve deeper into *time, making, say*, and *face*, suggesting tensions around growing up, decision-making, and emotional expression, especially in relation to paternal authority (*Papa*) and the daughters' restrained agency. Topic 5's inclusion of *veranda, went*, and *sit* conveys an almost static, observational existence—particularly of female characters like Uma and Mira-Masi—who dwell in the margins of action. Altogether, the modeling illustrates how Desai's novel navigates the **gendered construction of space, emotional suppression, and fragmented identity within a traditionally structured Indian family**, offering a poignant critique of cultural expectations imposed on women.

4. Topic Modeling of “The Calcutta Chromosome ” by Amitav Ghosh

Topic 1:	antar said man hand like urmila room face eyes open
Topic 2:	said know murugan urmila going way tell want time think
Topic 3:	said sonali urmila phulboni door room murugan went looked farley
Topic 4:	said murugan antar malaria like know ross cunningham time right
Topic 5:	murugan said antar urmila began mrs aratounian old ronnie head

The topic modeling analysis of *The Calcutta Chromosome* by Amitav Ghosh reveals a rich interplay between **science, mystery, and fragmented personal narratives**, reflecting the novel's postmodern structure and themes. Central to each topic are key figures like *Antar, Murugan, Urmila*, and *Ross*, pointing to the interconnected roles of **observation, research, and identity across time and place**. Topic 1 focuses on *man, hand, face*, and *eyes*, suggesting themes of physical embodiment and surveillance, reinforcing how bodies become sites of knowledge and experimentation. Topic 2, with terms like *going, way, tell*, and *think*, reflects a cognitive and narrative journey—Murugan's obsessive search for truth and the epistemological quest that drives the novel forward. Topic 3's inclusion of *Phulboni, door*, and *room* emphasizes spatial transitions and hidden knowledge, possibly tied to mystical or alternative histories. Topic 4 integrates *malaria, Ross*, and *Cunningham*, directly referencing the historical scientific backdrop of the novel, and brings out the tension between colonial research and indigenous knowledge systems. Topic 5 consolidates the presence of figures like *Mrs. Aratounian* and *Ronnie*, with words like *old, head*, and *began*, anchoring the narrative in institutional memory and generational shifts. Overall, these topics point to a text that negotiates between **gendered voices, colonial legacies, and subaltern agency**, using fragmented storytelling and layered character perspectives to deconstruct scientific authority and emphasize hidden, often feminized, modes of knowledge transmission.

- **Comparative study between the two novels based on Topic Modeling**

The topic modeling analysis of *Fasting, Feasting* by Anita Desai and *The Calcutta Chromosome* by Amitav Ghosh reveals distinct yet intersecting constructions of gendered experiences and emotional expressions within Indian English literature. In *Fasting, Feasting*, the topics are heavily centered on domestic and familial roles, with recurring references to characters like *Uma, Arun, Mama*, and *Masi*. These topics underscore the emotional labor and limited agency experienced by women, especially within the confines of traditional Indian households. The language of the topics—featuring verbs like "come," "went," "sit," and "said"—suggests a narrative of emotional entrapment and passive existence, particularly from the female perspective. In contrast, the topic modeling of *The Calcutta Chromosome* presents a more dynamic and fragmented narrative space, where characters such as *Antar, Murugan*, and *Urmila* move through a world of scientific mystery and epistemological inquiry. The presence of words

like "going," "know," "think," and "looked" highlights a more active engagement with the world, allowing for a broader and more nuanced portrayal of gender, especially through female figures who participate in unraveling hidden knowledge systems. Spatial references in Desai's novel tend to be confined to the home, symbolizing social restriction, while Ghosh's spatial terms suggest liminality and disruption, reflecting the novel's postcolonial and speculative dimensions. Ultimately, the contrast between the two novels through topic modeling demonstrates how Desai portrays emotional repression and gendered silence within familial and cultural expectations, whereas Ghosh offers a subversive and intellectually layered framework where gender intersects with science, power, and resistance. This comparative lens reveals how both authors engage with and critique cultural narratives through differing modes—Desai through intimate emotional interiority, and Ghosh through postcolonial subversion and mystery.

5. NER

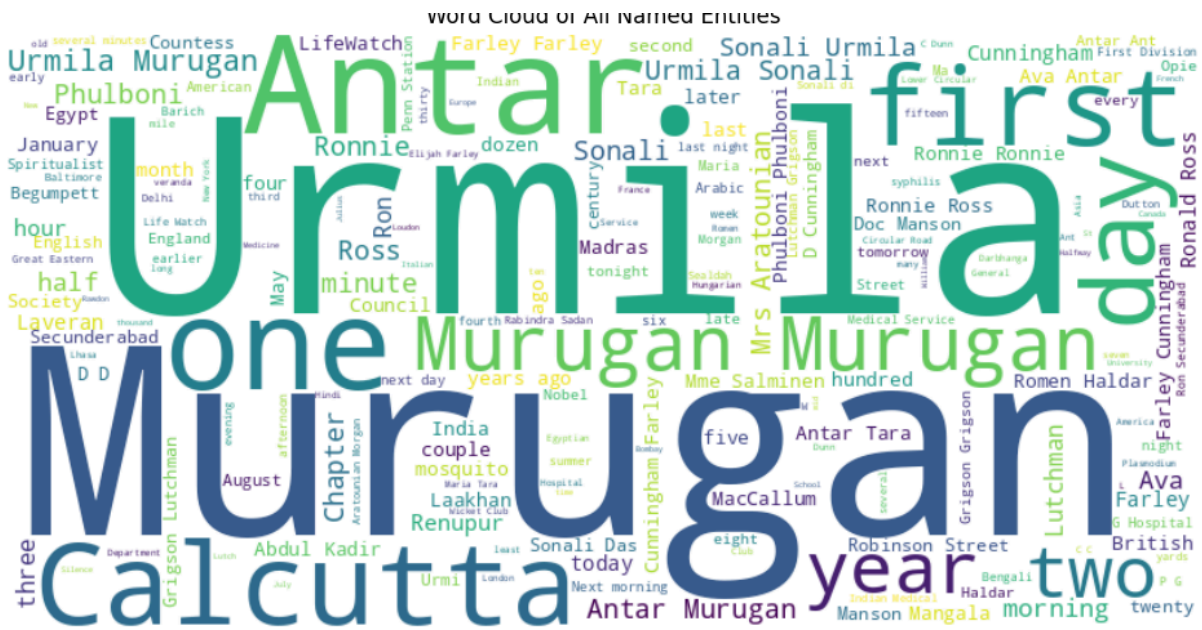
The NER output of *The Calcutta Chromosome* reveals a landscape dominated by named entities that underscore the novel's intricate engagement with science, identity, and postcolonial power structures. Entities such as **Murugan**, **Antar**, **Urmila**, **Phulboni**, **Ross**, and **Cunningham** anchor the narrative within a transnational and multi-layered historical context. The presence of both Indian and Western names highlights the novel's tension between colonial legacies and indigenous epistemologies. Importantly, the appearance of female characters like **Urmila** within these named entities suggests a deliberate positioning of women in roles of observation, intellect, and subtle resistance — contrasting with more traditional representations. The recurrence of institutional and locational entities (e.g., **stations**, **rooms**, **labs**) further grounds the text in spaces of research and movement, revealing how knowledge is gendered and spatially regulated. From a gendered perspective, the NER results show that while male figures often dominate the scientific discourse, female entities like Urmila act as connectors and questioners, navigating through emotional subtext and systemic silence. This highlights a layered construction of gendered experience — not through overt domesticity as in Desai's work, but through presence in epistemological inquiries and symbolic disruption. In sum, the NER findings reflect Ghosh's nuanced portrayal of gendered identity within a speculative, intellectual, and postcolonial matrix, emphasizing how cultural narratives are refracted through both visible authority and subtle emotional expression.

The NER output for *Fasting, Feasting* foregrounds a tightly confined world of family names, personal relationships, and domestic spaces, which reflects the novel's central concern with the emotional and social entrapment of women in traditional Indian households. Named entities such as **Uma**, **Arun**, **Mama**, **Papa**, and **Mira-masi** recur prominently, emphasizing familial hierarchies and the gendered roles each character performs. These names are rooted in intimacy and familiarity rather than public or institutional authority, highlighting how gender in Desai's novel is deeply tethered to emotional obligation, care work, and societal expectation. Spatial and temporal named entities—like references to **America** or **school**—tend to signify the contrast between aspiration and restriction, often drawn along gendered lines. For example, while Arun is allowed academic and physical mobility, Uma's world is marked by stasis, symbolized through repeated domestic and familial references. Moreover, the sparse appearance of institutional or professional named entities reinforces the limited public identity of female characters. In this way, the NER results reveal how emotional expressions are constructed through dependency, denial, and cultural silence, especially in the lives of women like Uma. Thus, Desai's text, as uncovered through NER, constructs a gendered world where emotional narratives are framed within confinement and unfulfilled desire, exposing cultural narratives of sacrifice, repression, and the invisibility of female subjectivity in Indian English literature.

- **Comparative Analysis of two novels based on NER**

In *Fasting, Feasting*, Anita Desai constructs gendered experience through emotional intimacy and domestic entrapment, focusing on characters like Uma, who embody the silenced, sacrificial roles of women within traditional Indian family systems. The NER results underscore this with recurring entities tied to familial roles (Mama, Papa, Uma, Arun) and confined spaces, reflecting a culture that assigns women emotional labor and denies them agency. Emotional expressions here are internalized — rooted in longing, repression, and invisibility. In contrast, Amitav Ghosh's *The Calcutta Chromosome* presents a more complex and decentralized construction of gender, where female characters like Urmila operate within intellectual and epistemological spaces. While male entities (Murugan, Antar) dominate the scientific discourse, women subtly subvert the patriarchal narrative by inhabiting spaces of mystery, silence, and alternative knowledge systems. Emotional expressions are not overt but encoded in the fabric of speculative storytelling and symbolic absence. The contrast lies in Desai's introspective, family-centered portrayal of gender and emotion versus Ghosh's postcolonial, intellectual treatment where emotion operates through enigma and erasure. Together, the two novels reflect the shifting cultural narratives in Indian English literature — from

2. “The Calcutta Chromosome ” by Amitav Ghosh



Limitations & Challenges

While this study offers new insights into gendered experiences and emotional expression in Indian English literature through digital humanities methods, several limitations and challenges must be acknowledged. Firstly, computational tools like Named Entity Recognition and topic modeling, though powerful, are primarily developed for general English corpora and may not fully capture the nuanced linguistic patterns or culturally specific expressions present in Indian English texts. Emotional and gendered language in these novels may be influenced by regional idioms, code-switching, and layered cultural meanings that automated tools might overlook or misclassify. Additionally, the relatively small dataset—focused on just two novels—limits the generalizability of the findings, making the study more exploratory than conclusive. There are also interpretive challenges in aligning algorithmically generated word clusters or topics with meaningful literary analysis, which requires careful contextual reading. Finally, the emerging nature of digital humanities in the Indian context means there are fewer comparative studies, models, and annotated corpora tailored for postcolonial and gender-focused research, posing constraints on methodological refinement and validation.

TakeAways

- The study reveals distinct ways gendered experiences are portrayed in Indian English literature, reflecting both traditional and evolving cultural narratives.
- In *Fasting, Feasting*, emotional expression is tied to domesticity, repression, and the feminine experience within rigid family roles.
- In *The Calcutta Chromosome*, gender and power are explored through mystery and science, emphasizing silenced voices and hidden agency, especially of female characters.
- Computational tools like NER, word clustering, and topic modeling helped uncover thematic patterns that might not be visible through close reading alone.
- Emotion and gender are not only themes but are linguistically embedded in how characters, actions, and spaces are described.
- The emotional vocabulary in Indian English literature often reflects cultural specificity, shaped by vernacular influence and local traditions.

The study shows the value of integrating digital humanities methods with literary and gender theory to uncover deeper patterns in postcolonial texts.

These findings contribute to the broader understanding of how Indian English novels negotiate identity, emotion, and gender within a socio-cultural framework.

Ai help for the Project

This research project was primarily self-driven, with around 80% of the coding and data visualization work completed independently by me. I handled tasks such as Named Entity Recognition, topic modeling, and word clustering using Python and NLP libraries like spaCy. I also created and interpreted visual diagrams to support the thematic analysis of the novels. While I engaged with existing literature and theoretical frameworks, the technical execution and integration of digital methods were managed on my own. This hands-on approach helped deepen both the analytical rigor and interdisciplinary scope of the study.

