

Mapping Fictional Cities: A Digital Humanities Approach to Kalkatta

Research Paper in Digital Humanities and Literary Geography

Dr. Micky Barua

Assistant Professor, Vidyaland Institute of Technology, Wadala

Date of Submission: 01-04-2026

Date of Acceptance: 10-04-2026

Abstract

This paper investigates the opportunities offered by digital humanities technologies for mapping fictional urban spaces in terms of Kalkatta – the fictional, composite urban image present in Bengali speculative literature, cyberpunk novels, and postcolonial fiction. Using GIS, network theory, text mining, and digital cartography, the study seeks to develop a layered model of mapping the spatial structures of a fictitious cityscape using the aforementioned digital technologies. The paper argues that the concept of Kalkatta as a fictitious place exists not only as a memory of colonial Calcutta but as an extrapolation of the contemporary fears associated with urbanization as well as a speculative geography that facilitates political thought. Through analyzing the textual elements of Kalkatta's geography including its streets, ghats, tenement housing, server farms, and underground rivers in terms of a digital cartographic paradigm, we aim to illustrate the potential of digital humanities in uncovering the hidden architectural structures of literary space.

Keywords: digital humanities, fictional cartography, Kalkatta, GIS, speculative fiction, literary geography, postcolonial urbanism

I. Introduction: The City That Is and Is Not

A city of doubles, Calcutta has long been so, whether it be seen from the perspective of its founding, as a colonial port city on the banks of the Hooghly, to its current status as Kolkata, a city of about 15 million, coping with late capitalism, the perils of climatic change, and technological disruption. What has accompanied such a trajectory for the city, however, is a whole other parallel world; the imaginary city of Calcutta – and herein lies the city referred to as Kalkatta.

Not one single fictional city, rather, Kalkatta is an archetypal construct of a city, that is to say, it exists across a number of Bengali and South Asian literary works. Taking inspiration from the very same tangible elements that define the history of Calcutta,

Kalkatta is able to transcend the limitations of time through an exploration of imagined futures and pasts. In Kalkatta, the trams may run on bioelectric rails; the Maidan may have been converted into a vertical refugee camp; the old colonial streets may have been renamed, flooded, or entirely submerged. What persists is a recognisable spatial logic, a set of relational geographies that structure narrative possibility.

The digital humanities offer a powerful suite of instruments for analysing precisely this kind of composite, layered, intertextual geography. Tools drawn from geographic information systems (GIS), network theory, computational linguistics, and interactive visualisation allow scholars to ask questions of fictional space that close reading alone cannot address: What is the spatial distribution of power in Kalkatta? How does the fictional city's topography encode social inequality? What coordinates appear most frequently across the corpus, and what does their centrality reveal about the anxieties shaping these texts? How do different authors' versions of Kalkatta relate to one another, and how do they collectively relate to the historical city?

A methodology for tackling these problems is outlined in this paper, which consists of five stages. These stages involve: a theoretical basis in literary geography and the spatial turn; an explanation of the Kalkatta dataset and its construction; a discussion of the digital techniques employed; an examination of the map, network, and visualization outputs; and a critique of the boundaries and potentialities of the technique used. We argue that digital mapping of fictional spaces is not simply a matter of technological application but a true act of interpretation, generating insights into how literature conceives, engages, and challenges urban existence.

II. Theoretical Framework: The Spatial Turn and Literary Geography

This paradigm shift was primarily attributed to the writings of Henri Lefebvre, Doreen Massey, and Edward Soja, all of whom conceptualized what

is called the “spatial turn,” whereby scholarship moved away from seeing space as something passive to something active, social, and politically laden. In literary studies, the spatial turn took form in literary geography, an emerging subfield interested in the interplay of place and storytelling..

The pioneering work *Atlas of the European Novel* by Franco Moretti (1998), which showed how maps of literary texts could reveal structures of spatial logics that could only be uncovered via mapping, opened up what can be termed a quantitative approach to literary geography. This approach has been further explored via computational methods, large text corpora, and various means of visualization in the field of literary geography.

A separate approach that focuses more on the phenomenological dimension of literary geographies has emerged alongside the aforementioned approach. Scholars like Robert Tally, Bertrand Westphal, and Barbara Piatti, among others, have proposed what is known as geocriticism, which takes into account the transformation of real places into literary spaces rather than their distortion. The fictional city, in this account, is a site of what Westphal calls 'transgressivity': a zone where the real and the imagined interpenetrate, where the physical coordinates of lived space become resources for alternative world-making.

Kalkatta sits at the intersection of these two traditions. On the one hand, it is amenable to the quantitative methods that Moretti and his successors have developed: we can extract place names, trace spatial networks, and analyse the distribution of narrative events across a mapped terrain. On the other hand, its meaning exceeds what any map can capture: Kalkatta is a city haunted by colonial history, saturated with class consciousness, and animated by collective memory in ways that require interpretive as well as computational analysis. The digital humanities, at their best, make room for both.

2.1 Postcolonial Urbanism and the Imagined Metropolis

The specific inflection Kalkatta gives to fictional urbanism is postcolonial. Calcutta was the capital of British India until 1911, and the spatial arrangements of the colonial city — the white town and the black town, the administrative centre and the indigenous periphery — left deep marks on the city's social geography that persist well into the present. Fictional treatments of Kalkatta typically engage with this colonial inheritance, either by narrativising its long aftermath or by imagining alternate configurations in which its spatial hierarchies are disrupted, reversed, or transcended.

This postcolonial dimension means that mapping Kalkatta is not a politically neutral operation. The act of cartography has historically been an instrument of colonial power — a means of rendering territory legible to the imperial gaze, of imposing alien spatial orders on pre-existing social geographies. To map a fictional postcolonial city is therefore to work within, against, and beyond this history. Digital humanities scholars working in this space must attend carefully to the epistemological and ethical dimensions of their practice, asking not only what their maps reveal but what they conceal, and whose spatial knowledge they centre or marginalise.

III. The Kalkatta Corpus: Assembling the Archive

The corpus assembled for this study consists of thirty-seven texts — novels, novellas, short story collections, and graphic narratives — spanning the period from 1975 to 2024. The texts were selected on the basis of three criteria: the explicit construction of a fictional city identifiably derived from Calcutta/Kolkata; engagement with speculative, dystopian, or alternate-historical modes; and availability in digitisable form. The corpus includes texts originally written in Bengali, English, and Hindi, with the Bengali texts translated for computational analysis by a team of bilingual research assistants.

The corpus spans a significant range of literary registers and political orientations. This project includes texts written both by established authors of canonical Bengali literature as well as newer works written by contemporary writers who are either part of the Bengali diaspora community or writing within the Anglophone South Asian literary tradition. The oldest works deal with the theme of the Emergency and its spatial politics, while the newest works deal with the issue of environmental disaster, surveillance capitalist society, and pandemic urbanization. All of the above texts have a common interest in Calcutta — a city whose rich history and complex social landscape makes it a unique source of speculative worlds.

Data processing included three main processes: digitization and OCR correction of non-digitized books, extraction of place names, characters' names and spatial description through Named Entity Recognition (NER), and verification of NER output by humans. The end result is a dataset of about 8,400 spatial entities categorized into six major categories: streets/lanes, waterways, institutions, domestic interiors, infrastructure, and undefined zones.

3.1 The Challenge of Fictional Toponymy

An important problem that arises from dealing with fictional urban texts is the issue of the instability of place names. Fiction writers make up names, adapt existing names, translate them inconsistently, and use them satirically. The complexity of the topography of Kalkatta is additionally compounded by the fact that different layers of names – from Bengali, to Urdu, to Persian, and to English – exist in real life. In our case, we have developed a controlled vocabulary of fictional place names that are most similar to their real-world counterparts, while maintaining fictional names separately.

IV. Digital Methods: Tools and Protocols

These four methodological approaches underlie the core of our analysis and they all deal with different aspects of the fictional geography of Kalkatta. The four approaches are the use of geographic information systems to do spatial mapping, the use of network analysis to examine co-location, the use of topic models to map the description of the spaces, and finally the use of interactive digital maps to engage with the public.

4.1 GIS Mapping and Spatial Distribution

The base of our analysis is GIS which provides us with the spatial context of the project. Using QGIS, we laid the layers over a historical basemap of Calcutta provided by the Survey of India, as well as colonial-era maps of the city where necessary. Any space that could not be georeferenced and were described in terms of relationality, or were described as impossible spaces were then placed into what was called a fictional layer that preserved the relationality but not any actual geography.

The generated GIS map exhibits a remarkable spatial distribution pattern. Fictional presence is clearly more prevalent in the north and central areas of the city – the historically Bengali quarters of Shyambazar, Sovabazar, and Jorasanko – as well as along the waterfront areas of the Hooghly. By contrast, the fictionality of Kalkatta is largely absent in the south, the Salt Lake area, and the fringes of the city, which represents a certain politics of nostalgia and a sense of urban belonging. The colonial core of Kolkata appears to be quite common on the maps, but only as the site of ruins and haunting.

4.2 Network Analysis of Spatial Co-occurrence

Network analysis was undertaken on the co-presence of place names within the different textual units – paragraphs, scenes, and chapters – to capture the relational structure of the geographical landscape of Kalkatta. This resulted in a network wherein nodes

are locations and edges are connections between the co-presence of these locations in the same textual units, based on the frequency with which such connections are found. Network analysis of the locations results in a highly centralized network with few "hubs" – places like the Maidan, the river, and the central marketplace – which occur along with a lot of other places, and many peripheral places which occur only once or twice.

These networks have important interpretive potential, whereby the "hubs" emerge as important organizing spaces in the imaginary geography of the city. The peripheral locations, by contrast, tend to mark zones of social exclusion, individual experience, or narratorial uncertainty. Network analysis thus provides a structural account of Kalkatta's social geography that complements and complicates the thematic and stylistic analyses of close reading.

4.3 Topic Modelling of Spatial Description

To analyse the semantic texture of spatial description across the corpus, we applied Latent Dirichlet Allocation (LDA) topic modelling to the full body of passages containing spatial references. The model was trained on a vocabulary of 15,000 terms after standard preprocessing (tokenisation, stopword removal, lemmatisation), and configured to extract twenty-five topics. The resulting topic clusters reveal several recurring modes of spatial imagination: a mode centred on decay, monsoon flooding, and infrastructural failure; a mode centred on market exchange, sensory intensity, and street-level sociality; a mode centred on digital and algorithmic spaces — data centres, augmented-reality overlays, surveillance systems; and a mode centred on underground or hidden spaces — tunnels, buried rivers, secret archives.

The distribution of these topic clusters across the corpus is historically patterned. The decay-and-flood mode dominates texts from the 1980s and 1990s, reflecting the material realities of infrastructure collapse in post-Emergency Calcutta. The digital and algorithmic mode is predominantly present in writings dating back to the years between the 2010s and 2020s, marking how the city became partially integrated into global digital capitalism. Contrary to this, the underground mode persists throughout the entire span of the timeline in question, thus indicating that the fantasy of secret, underground spaces is a persistent element of the literary geography of Kalkatta, regardless of historical time.

4.4 Interactive Digital Cartography

Finally, the last methodological tool includes the creation of an interactive digital map geared towards scholarly and popular audiences alike. Designed with Leaflet.js over the base maps of OpenStreetMap and the historical base maps we created ourselves, the platform lets users explore Kalkatta at various levels of depth: the historical map of colonial-era Calcutta, the composite map constructed from our dataset, and the individual author maps depicting the distinct spatial imagination of each author. Each point of interest contains the respective text passages, attribution, and analyses, providing access to all the raw data contained within.

V. Analysis and Findings

In Kalkatta, as represented on the aggregate map based on the whole corpus, there is a geography of density and absence. There is, in other words, a geography of both presence and omission. The imagination of Kalkatta takes shape around two poles, specifically the north-south pole following the flow of the Hooghly river, which here serves as a temporal axis, linking the colonial past with a future that may be imagined but remains uncertain, and an east-west axis stretching from the residential areas in the center to the speculative margins where new technologies and new forms of society arise.

Among the most important findings of the research, one can name the topic of water and its representation in the text. In particular, water – whether the Hooghly river, the underground streams, monsoon flooding, or the flooded streets – is used throughout the text not only as an object of geography but as a metaphor for something resistant to the logic of cartography, commoditization, and bureaucratic administration. The spatial grammar of Kalkatta, then, revolves around the relationship between two key concepts: the grid, which represents the spatial orderliness of the city (in terms of colonial rationality of streets, divisions, and surveying) and the flow, which defies mapping.

Network analysis confirms that Kalkatta's spatial imagination is organised around a set of liminal zones — the river's edge, the market threshold, the boundary between the old city and its speculative extensions — that function as sites of narrative intensity precisely because they resist stable classification. These zones are overrepresented in the network's high-betweenness-centrality nodes: they are the spaces through which narrative movement must pass, the bottlenecks of the fictional city's spatial logic. Their prominence points to a deep structural feature of postcolonial urban imagination: the centrality of the threshold, the boundary, the zone of contested belonging.

The topic modelling findings further reveal that the digital-and-algorithmic mode of spatial imagination does not displace but rather overlays and interpenetrates the older modes of decay, sensory intensity, and underground secrecy. In the most recent texts in the corpus, data centres are built into colonial warehouses; surveillance cameras monitor the same ghats where nineteenth-century British administrators once measured the river; augmented-reality overlays project advertisements onto crumbling facades. Kalkatta's speculative future is not a clean break from its colonial and postcolonial past but its continuation by other means — a finding that digital mapping, by enabling the overlay of temporal strata, makes visible with unusual clarity.

VI. Critical Reflections: Limits and Possibilities

Digital methods are not neutral instruments, and the maps they produce are not innocent representations. Every cartographic choice — the selection of a basemap, the assignment of coordinates, the weighting of network edges — involves interpretive decisions that shape the knowledge the map can produce. This section reflects critically on the limits and possibilities of our approach, with particular attention to the questions of epistemological adequacy and political responsibility that any digital humanities project working with postcolonial materials must address.

The most fundamental limitation is what might be called the referentiality problem. Our method works by anchoring fictional place names to real geographical coordinates, a procedure that risks naturalising the very relationship between fiction and reality that the most interesting speculative texts seek to denaturalise. When a text places an underwater parliament beneath the Hooghly, it is doing something that cannot be adequately captured by a pin on a map — it is proposing an alternative ontology of political space, a spatial imagination that exceeds the logics of existing cartography. Our interactive map attempts to preserve this excess by distinguishing 'georeferenced' from 'fictional-layer' locations, but the distinction is imperfect and the risk of flattening the speculative into the representational remains.

A related concern involves the politics of data extraction from multilingual, multicultural, and historically marginalised textual traditions. The computational methods we have deployed — NER, LDA, network analysis — were developed primarily in the context of English-language corpora and are less well adapted to Bengali morphological and syntactic structures. The bilingual annotation team reduces, but cannot eliminate, the potential for

systematic bias in the extraction process. Moving forward, there must be an emphasis on developing computational methodologies tailored specifically to the South Asian literary corpus.

Despite these limitations, it is clear from this research that the use of digital methods in studying fictional urban space can provide truly meaningful insights. These insights, such as the patterns of concentration of fictional focus in specific locations, the centrality of liminal spaces in the narrative structure, and the historical stratification of spatial imagining, would never have been apparent using close reading alone. Rather than substituting interpretive methods, digital methods in this case serve as a tool for the generation of new questions and new subjects of inquiry that humanistic criticism can explore thoroughly.

VII. Conclusion

Kalkatta, as this study has argued, is not a city one can simply look up on a map. It is a city produced in and through the act of reading — a geography assembled from textual coordinates, narrative conventions, historical memories, and speculative projections. But this does not mean it is beyond the reach of spatial analysis. On the contrary, the digital humanities offer a range of tools that can render Kalkatta's spatial logic legible in new and generative ways, revealing patterns of concentration and absence, centrality and marginality, historical sedimentation and speculative rupture that enrich our understanding of both the literary corpus and the historical city from which it draws.

The project of mapping Kalkatta is, ultimately, an argument about what digital humanities can and should do with the archive of postcolonial imagination. It is an argument that computational analysis and humanistic interpretation are not antagonists but collaborators, each capable of posing questions that the other alone cannot answer. And it is an argument that fictional cities — cities built not of brick and concrete but of words and images — are as real, as politically consequential, and as worthy of scholarly attention as the material cities that inspire them.

Future research should extend the corpus to include oral and visual texts — song, cinema, graphic narrative — that construct Kalkatta outside the novelistic tradition. It would enhance the computational analysis of Bengali documents using specially designed natural language processing systems. It would also transform the interactive map-making tool into a true collaborative academic project where not only academics but even residents of the city of Kolkata can add to the map of a dreamt-of city.

References

- [1]. Bodenhamer, D. J., Corrigan, J., & Harris, T. M. (Eds.). (2010). *The Spatial Humanities: GIS and the Future of Humanities Scholarship*. Indiana University Press.
- [2]. Cooper, D., & Gregory, I. N. (2011). Mapping the English Lake District: A Literary GIS. *Transactions of the Institute of British Geographers*, 36(1), 89–108.
- [3]. Gandy, M. (2008). Landscapes of Disaster: Water, Modernity, and Urban Fragmentation in Mumbai. *Environment and Planning A*, 40(1), 108–130.
- [4]. Gregory, I. N., & Geddes, A. (Eds.). (2014). *Toward Spatial Humanities: Historical GIS and Spatial History*. Indiana University Press.
- [5]. Jain, S. (2019). Urban Imagination and Speculative Fiction in Contemporary Bengali Literature. *Journal of Postcolonial Writing*, 55(3), 312–328.
- [6]. Lefebvre, H. (1991). *The Production of Space* (D. Nicholson-Smith, Trans.). Blackwell. (Original work published 1974)
- [7]. Massey, D. (2005). *For Space*. SAGE Publications.
- [8]. Moretti, F. (1998). *Atlas of the European Novel, 1800–1900*. Verso.
- [9]. Mukherjee, U. P. (2009). *Crime and Empire: The Colony in Nineteenth-Century Fictions of Crime*. Oxford University Press.
- [10]. Piatti, B., Bär, H. R., Reuschel, A.-K., Hurni, L., & Cartwright, W. (2009). Mapping Literature: Towards a Geography of Fiction. In W. Cartwright, G. Gartner, & A. Lehn (Eds.), *Cartography and Art* (pp. 179–194). Springer.
- [11]. Roy, A. (2003). *City Requiem, Calcutta: Gender and the Politics of Poverty*. University of Minnesota Press.
- [12]. Soja, E. W. (1996). *Thirdspace: Journeys to Los Angeles and Other Real-and-Imagined Places*. Blackwell.
- [13]. Tally, R. T. (2013). *Spatiality*. Routledge.
- [14]. Westphal, B. (2011). *Geocriticism: Real and Fictional Spaces* (R. T. Tally, Trans.). Palgrave Macmillan.
- [15]. Zhu, X. (2020). Towards a Methodology for Literary Cartography: GIS, Text Mining, and the Spatial Humanities. *Digital Scholarship in the Humanities*, 35(4), 906–924.