

‘Multiplicity Embarrasses the Eye’: the Digital Mapping of Literary Edinburgh

James Loxley, Beatrice Alex, Miranda Anderson, Uta Hinrichs, Claire Grover, David Harris-Birtill, Tara Thomson, Aaron Quigley, and Jon Oberlander (Universities of Edinburgh and St Andrews)

Literature, Geocriticism, and Digital Opportunities

There has long been a deep intertwining of the concepts, experiences and memories of space and place and the practice of literature. But while it might be possible to conduct a spatially-aware critical analysis of any text, not all texts manifest such self-conscious spatial awareness. Literary history, nevertheless, is rich with genres and modes that do, from classical pastoral and renaissance estate verse to metadrama, utopian writing and concrete poetry. Any critic seeking to give an analytical account of such writing will find it hard to overlook its immersive engagement with space and spatiality. And literary criticism has also enriched its vocabulary with concepts drawn from a range of modern and contemporary theoretical sources. Mikhail Bakhtin’s notion of the chronotope, he suggested, named ‘a formally constitutive category of literature’, marking an irreducible spatiality (intertwined, for Bakhtin, with temporality) which was yet manifest only in particular, historically and generically specific, configurations.¹ Michel de Certeau developed a very different approach, distinguishing space from place in a distinctive fashion that has had a significant influence on literary critics.² Yu Fu Tuan’s own conceptualisation of this distinction has also had a literary critical resonance.³ Henri Lefebvre’s account of the ‘production of space’, with its conceptual triad of ‘spatial practice’, ‘representation of space’ and ‘representational space’, has also exerted its pull.⁴

Recently, however, literary critical interest in spatiality has reached a new intensity. No single label, let alone conceptual orientation, covers the range and kinds of work being undertaken. Where practitioners and observers have attempted definition, they have of necessity had to think

inclusively. So Robert Tally, one of the most influential current critics working in this area, has sought to cast the net widely. As he says:

I consider spatial literary studies—whether doing business as geocriticism, literary geography, the spatial humanities, or using some other moniker—as a multiform critical practice that would include almost any approach to the text that focuses attention on space, place, or mapping, whether within the confines of the text, in reference to the outside world, or some combination of the two. What Edward W. Soja has termed the real-and-imagined places of literature, criticism, history, and theory, as well as of our own abstract conceptions and lived experience, these constitute the practical domain for spatial literary studies.⁵

For Tally, geocriticism itself names ‘something like a general comportment toward the text, rather than a discrete methodology with its own set of rules and conventions’.⁶ While his is not the only possible understanding of what either ‘geocriticism’ or ‘spatial literary studies’ might be, it is one that manages to grasp both the centrifugal and centripetal forces at work in current research in the field. And while Tally is obviously alert to the great differences between – to cite his examples – ‘using geographical information systems (G.I.S.) to chart a novel's character or plot trajectories along the physical topography of a given region’ and ‘examining the concepts of *deterritorialization* and *reterritorialization* in theorizing matters of poststructuralist geophilosophy’ he is right to seek to maintain a dialogue between them.

Central to the geocritical approach taken by Tally and others (particularly Bertrand Westphal, some of whose work Tally has translated into English⁷) is the notion of mapping. Literary texts have a cartographic dimension, to which a geocritical approach is attuned and which it seeks to bring out. Here, geocriticism perhaps insistently literalises what has been a sometimes metaphorical approach to cartography as a way of accounting for, or making sense of, a literary text's spatiality. This is an approach that chimes with long term tendencies in the influential critical work of Franco

Moretti. Moretti's own sense of the explanatory power of cartographic representation, first thoroughly evident in his *Atlas of the European Novel, 1800-1900*, has provoked and complemented his interest in the usefulness of computational methods in literary analysis, the practice of 'distant reading' with which his name is now definitively associated.⁸ The convergence of the cartographic and the computational demonstrated by Moretti is part and parcel of a broader interest in such possibilities, visible in different ways in Ian Gregory and David Cooper's *Mapping the Lakes*, Barbara Piatti's *Literary Atlas of Europe*, Shelley Fisher Fishkin's collaborative *Digital Palimpsest Mapping Project*, and the 'z-axis' work of the *Modernist Versions Project*, to name only four recent examples.⁹

It is possible, too, that digital approaches might offer solutions to geocritical problems which are insoluble in other ways. Tally himself has raised what, from one perspective, is an insuperable limit to Westphal's project:

How does one determine exactly which texts could, in the aggregate, reasonably constitute a meaningful body of material with which to analyze the literary representations of a given geographical site? ... With certain cities, such as Paris, London, Rome, or New York, the almost mythic status of these places and the seemingly innumerable textual references to them render any geocritical analysis, at least those laying claim to a kind of scientific value, impossible. As Westphal admits, 'to attempt to undertake a full-scale geocritical analysis of those hotspots would be madness.' A geo-centered method, if it aims truly to avoid the perception of bias, seems somewhat doomed from the start.¹⁰

There is here an echo of the modern and postmodern experience of the cityscape – a space or place too densely and variously populated to map or hold in a single prospective view.

But perhaps peeking beyond the horizon of possibility imagined here need not be 'madness' after all. Text mining and data visualisation technologies open up the possibility of reading very large corpora for their geocritical significance, or at least of creating the conditions for more

thoroughgoing geocritical exploration to take place. This is a way of thinking about the relationship between literature and cartography at several different levels – not just the ‘internal’ map of a literary work, or the literary use of non-literary locations, but also the ways in which multiple texts might meaningfully coincide or vary in the way that they engage with the same loci. Where smaller scale studies might focus most on the former possibilities, the larger scale aggregation and alignment of literary representations of place offers opportunities for interaction and analysis that might well lead us to wonder about what ‘full-scale geocritical analysis’, and the attendant qualifying notion of ‘scientific value’, can actually mean.

The Palimpsest Project: Mapping Literary Edinburgh

The Palimpsest project, undertaken at the universities of Edinburgh and St Andrews in 2014-15, provides an example of this larger scale digital analysis. The project was conducted by an interdisciplinary research team of literature and informatics scholars, including specialists in text mining, geoparsing, data visualisation, and computer-human interaction. Taking up the challenges posed by Tally, the Palimpsest team sought to devise a way in which one could digitally map narrative representations of a city, in this case Edinburgh, translating a literary critical approach to place into a digitally intelligible model.

The self-conscious ‘literariness’ of Edinburgh, and its existing history of literary tourism, make it an ideal case for this type of research. It stands rather overtly as one of those ‘real-and-imagined places of literature’ that Tally describes:¹¹ a city that has been extensively reimagined through narrative and prides itself as a literary locale. In the 1840s, a 60-metre tall monument to Sir Walter Scott was erected, which now looms over the city’s central train station, itself named after one of Scott’s novels. Edinburgh was already regarded by the 19th century as a city defined by, and imagined through, its literary spaces. Laurence Hutton’s *Literary Landmarks of Edinburgh* begins with the claim: ‘No city in the world of its age and size – for Athens is older and London is larger – is so

rich as Edinburgh in its literary associations, and no citizens anywhere show so much respect and so much fondness for the history and traditions of their literary men'.¹² He continues to marvel that this interest is not just found in the educated classes, but 'among the more poorly housed and the less educated classes, in whom one would least expect to find it'.¹³ His opening portrait of Edinburgh is a city in which ordinary, everyday life is suffused with literary memory, even through its dirtiest pubs and its smallest wynds. This literary spatial memory stretches in Hutton's narrative from the 17th century onward, beginning with William Drummond's 'Hie Schule', built in 1567, and leading to Dr John Brown's mid-19th century Rutland Street home.

Hutton's approach to the city is still reflected in Edinburgh literary tourism. Visitors can take literary tours, walking past the homes of great authors such as Robert Burns or Arthur Conan Doyle, sitting in pubs visited by Johnson and Boswell, or stopping at the cafés in which J.K. Rowling wrote the first Harry Potter novel. Most of these tours approach Edinburgh as the real place that nurtured and inspired great works of literature, although more recently some take visitors past Edinburgh locations that are imagined in literary works, such as the tours of Irvine Welsh's Leith or locations that feature in Ian Rankin's Rebus novels. In 2004, Edinburgh was the first city to win a UNESCO World City of Literature designation, and while its literary tourist industry continues to grow, the focus of most tours remains on landmarks and topographical routes, including limited literary content.

The Palimpsest project wanted to instead offer a different experience of literary Edinburgh – a more text-centered experience, that would enable users to explore narratives from different periods clustered at one location, or narratives that span locations across the city. Building on findings from a prototype developed in 2012 from an idea by Miranda Anderson, the project team aimed to gather a more expansive dataset by mining several large collections of digitised text for works set in Edinburgh. We envisioned two visualisation tools through which end users could explore our database: an interactive map of Edinburgh, featuring narrative extracts set in locations around the city, and a complementary mobile application that would allow users to move through

narrative imaginings of the city *in situ*. The database interface and suite of visualisations was released online under the name *Lit Long: Edinburgh*.¹⁴ Rather than using off-the-shelf software, the team developed these tools specifically for the Palimpsest project; however, many of their features draw on familiar visualisations, such as Google Maps, GPS navigation, word clouds, and histograms.

The question of which texts would form the dataset for *Lit Long*, the ‘meaningful body of material’ to which Tally points, was left, at least initially, to the algorithmic processes described in section three.¹⁵ However, human curation was required to ensure the selected texts met our literary critical criteria, and these criteria were devised prior to the text mining, in order to guide the computing processes. We were looking for ‘Edinburgh-centric’ works, which either belonged to a recognisably literary genre, such as the novel or short story, or had strong narrative or locodescriptive components – both memoirs and travel journals, for example, were included. We chose the use and frequency of place names as our primary marker in defining an ‘Edinburgh-centric’ work, and our gazetteer of Edinburgh place names guided our text searches, geolocation processes, and shaped the end visualisations.

Text Analysis and the Geo-location of Literary Place

The text processing work carried out for Palimpsest involved three main steps: (i) the creation of an Edinburgh gazetteer; (ii) retrieval and ranking of Edinburgh-specific literature candidates, and (iii) textual analysis of the Edinburgh-centric literature. Each of these tasks presented conceptual and practical challenges. There is no freely available gazetteer which includes all the Edinburgh place names the use of which we wanted to capture. These place names have different granularity, from area names (*Portobello, Cramond*), through street names (*The Royal Mile, Cockburn Street*) to open spaces (*The Meadows, Princes Street Gardens*), buildings (*Craigmillar Castle, Holyrood Palace*), statues and monuments (*Greyfriars Bobby, The Scott Monument*) etc. Therefore, a prerequisite for georeferencing was to create an Edinburgh gazetteer by aggregating

information from a variety of different sources. For street names we used OS Locator (OSL), for building-level information we used the Canmore site records database from the Royal Commission on the Ancient and Historic Monuments of Scotland (RCAHMS). For other information from area names through to pub names, we used an Edinburgh subset of Open Street Map (OSM). The aim was to create a gazetteer which could be used both as a place name lexicon when identifying potential place names in text and as a gazetteer for georeferencing, i.e. assigning coordinates to recognised place names.

The aggregation process involved converting records from the three sources into a common XML format followed by a data clean-up stage to discard records which might trigger faulty recognition of place names in text. For example, RCAHMS has records for places with generic names such as *Station House* or *Barracks*, as well as records for residential houses with names such as *Bonny Views*. OSM has records for numerous modern day businesses such as *Bay of Bengal* (a restaurant) and *Blossom* (a guest house). We attempted to exclude records such as these semi-automatically but the final gazetteer still contains many questionable Edinburgh place names, e.g. *Alien Rock* (a climbing centre), *Alpine Garden* (part of the Royal Botanic Gardens Edinburgh), *Beach House* (generic descriptor), and *The Waiting Room* (a pub). The presence of this kind of record in the gazetteer, however, does not seem to have had too deleterious an effect on place name recognition and subsequent georeferencing. There are some place names which occur in the Palimpsest books for which none of the three sources has a record. These are mostly historical forms of modern place names or spelling variants (for example, *Cowgate-port*, *Nor' Loch*, *Edinboro*). For cases where such an omission has been observed, we have manually added appropriate records. The final version of the Edinburgh gazetteer contains 13,064 records corresponding to 10,204 unique place names.

The aim of the retrieval and ranking component is to identify in the pool of accessible literature those works which are likely candidates for being Edinburgh-specific and rank them. The goal was to uncover items of literature which would not immediately come to mind when being asked to name literature set in Edinburgh. We processed five major literary collections containing

over 388,000 books in total, including world public domain works from HathiTrust, the British Library Nineteenth Century Books collection, out-of-copyright Project Gutenberg books, the Oxford Text Archive data and works from the National Library of Scotland.

In order to retrieve candidates of Edinburgh-specific literature, all of the literary data was indexed using Indri 5.6 and ranked using a set of 1,633 Edinburgh place name queries. We used the Indri inference network language model based ranking approach.¹⁶ In our case, the ranking score of a document is computed by combining the score for the location query retrieved from the content of a book with a score based on information in the metadata of the book. For example, the ranking is increased given the presence of a set of favoured Library of Congress codes (including PR, DA, PZ, PN and PS) as well as given a list of relevant subject terms (edinburgh, scotland, literature, fiction, novel, poetry, poem, story, stories, drama, novella, english, biography, ballads, ballad, scottish). This is to allow documents with such metadata information to appear higher in the ranking. At the same time, the ranking score is down-weighted for ambiguity of Edinburgh place names in order to push documents which mention place names most likely not referring to a location within Edinburgh, like *Trinity*, down the list. The output of the document retrieval component is a set of ranked Edinburgh-specific candidate documents per collection. The ranked output was visualised using a curation tool developed by the SACHI group at the University of St Andrews. The curation tool was used for manual curation in order to guarantee that the final Palimpsest data set is of high recall as well as of high precision.

During the project, we fine-tuned the ranking component based on feedback from the curators after a two-week pilot curation phase. We considered additional ambiguous place names identified by the curators, removed documents containing non-literary title words (e.g. dictionary, catalogue, etc.) and ensured that the place name *Edinburgh* or one of its variants (e.g. *Edinburrie*, *Embra*, etc.) occurred at least once in the document. Alex et al. showed that these three measures lead to a small improvement in the mean average precision (MAP) of the ranking on a test set created during the pilot.¹⁷ Most importantly, however, they resulted in a large reduction in the

number of ranked documents (60%) needing curation which considerably reduced the workload of the curators. Since we conducted this work, Ted Underwood has published a methodology and code for identifying genre within the HathiTrust data.¹⁸ Information on the genre (if not already available in the metadata) would considerably improve the quality of the ranked output as much of the manual curation of the ranked output was to differentiate between fictional and non-fictional works containing Edinburgh-based locations.

Using the optimised retrieval and ranking component, a total of 33,277 ranked documents were presented to the curators. Documents were ranked per collection as the type of metadata information available differed considerably. As a result of time constraints and evaluating the candidates further down the list in the ranked output, the decision was made only to curate the top 10% per collection manually. This resulted in 503 out-of-copyright documents which were considered to be part of the Palimpsest corpus and therefore contained an Edinburgh setting. A further 43 documents from modern authors for which we received permission from publishers were also added to this corpus.

The text mining part of the system is done using a Palimpsest-adapted version of the Edinburgh Geoparser.¹⁹ This contains two main components, a text mining pipeline for recognising place names and other entities in text and a geographic ambiguity resolution component which chooses between competing interpretations of place names (i.e. competing geographic coordinates) given their textual context.

The text mining pipeline first converts an input text into a common XML format and then each stage of processing incrementally adds annotations into the mark-up. First the text is segmented into paragraphs which are tokenised to add word and sentence elements. Words are part-of-speech tagged and lemmatised and then Named Entity Recognition (NER) is performed using hand-written rule sets combined with lexical look-up. For place name recognition, extensive lexicons of place names both from the UK and the rest of the world are used, and in the Palimpsest system

this stage is augmented to include a lexicon of Edinburgh place names derived from the Edinburgh Gazetteer.

The output of the text mining contains named entity annotations for dates, person names and place names. This is input to the geo-resolution step which looks up the place names in one or more gazetteers and ranks the results to arrive at the most probable interpretation given the context of the document. In Palimpsest, look-up in the Edinburgh gazetteer is done before look-up in more general Ordnance Survey and Geonames gazetteers which the system accesses via Edina's Unlock Places web service. Ranking uses heuristics combined with weighting of information such as geographic feature and size. We assume that a degree of geographic coherence holds within documents in that the relevant text is more likely to mention many places in a single area rather than a set of geographically unrelated places. To model this, proximity between gazetteer records for all the places mentioned in the document is strongly weighted to ensure that all locations mutually constrain one another to be as close together as possible. Thus the highest ranked interpretation of *Haymarket* will be the one in Edinburgh in a document that contains many Edinburgh or Scottish place names rather than the one in London in a document with more London or English place names.

The results of geo-resolution are added as XML annotations in the document and all that remains is to mark-up the immediate context of each place name for display in the Palimpsest interfaces. We call the context surrounding a geo-referenced Edinburgh place name mention a Palimpsest snippet. In the final system implementation, we set this context to be the sentence containing the location as well as the previous and following sentence without crossing paragraph boundaries.

The Palimpsest snippets were also ranked by an “interestingness” score (i-score). This was inspired by work on automatic prediction of text aesthetics and interestingness.²⁰ The aim is to rank snippets per document to give those snippets where the Edinburgh place name is not just a mention in passing more importance and therefore make them appear earlier on in the user interfaces. We

compute this score by checking for a number of features, including snippet length, the presence of multiple Edinburgh-based locations in snippet, the presence of at least one Edinburgh-based location (excluding variants of Edinburgh), an adjective or adverb appearing in the snippet, the presence of different forms of certain verbs (be, do, say or go) and word repetition within the snippet. The i-score is computed by treating each of the features equally and can range between a value of 0 and 1 where 1 represents snippets for which all features apply and 0 those where none of the features apply.

The output of the text processing components is fed into a database which provides the input data for the user interfaces. It comprises of over 550 literary works mentioning 1,600 unique Edinburgh locations mentioned in more than 47,000 literary excerpts. This is an especially rich, large scale dataset for a geocritical project, and constitutes a start, at least, on the work of gaining an effective mapping of the kind of literary 'high place' mentioned by Westphal and Tally.

Visualising Literary Place

The accessibility and utility of such a dataset, however, depends on its user interfaces. The study of geographic space and place from a literary perspective seems to automatically call for some sort of visual tools - call them maps - that facilitate explorations as well as the communication of patterns. Canonically, geographic locations are mapped to 2D space based on their unique latitude and longitude in order to explore or analyse spatial relations that would be otherwise impossible or, at least, difficult to decipher. Geospatial data is unique in that it comes with an inherent position which can be utilized in visual representations. In contrast, data extracted from literary texts such as the characteristics of the chosen language, applied grammatical structures, or more content-related themes and relations between literary characters are more abstract – a range of different types of visual mappings may be plausible to visualize this sort of data. For instance, a group of literary characters could be represented as dots positioned in a single large circle with connecting lines

representing their relations. When dealing with place names in a literary text one often intuitively thinks of these locations in terms of their geospatial meaning. Visualising literary space is challenged by this notion of abstract text on the one hand, and inherent geospatial meanings that invite conventional geospatial mapping on the other.

The visualisations and interactive interfaces we designed as part of the Palimpsest project - which were published online under the more easily pronounceable title of 'LitLong' - are based on the idea of making the meaning of place in Edinburgh literature explorable for different audiences and in different contexts, once again reflecting the sheer multiplicity of literary interest and experience provoked by the city. Researchers, for example, may want to apply an analytical perspective to explore the data in-depth from different perspectives focusing on authors, titles, themes and genre, publication year, or particular place-names mentions as an entry point. In contrast, a general-interest audience may want to explore literature related to selected books or well-known neighbourhoods in Edinburgh. General-interest audiences may include visitors who wish to explore how famous places such as the Royal Mile have been mentioned in literature; people living in the city may experience known and familiar locations in a new way and with fresh eyes by exploring related literary work – as Renton describes it in *Trainspotting*: 'They say you have to live in a place to know it, but you have to come fresh tae it tae really see it.'²¹ Interviews we conducted with literary scholars early on in the project indicated that in-situ explorations are of interest to academic scholars who may want to experience the places mentioned in literature themselves to facilitate interpretation.

To address these different potential exploration scenarios we developed two different visual interfaces that can be considered a first approach to explore this design space of visualizing the literary layer of a city. The Location Visualiser is a visual web-interface that allows the exploration of literary snippets related to places in Edinburgh from a birds-eye point of view. The LitLong App is a mobile application that runs on a smart phone, and shows corresponding literary snippets as people explore Edinburgh in-situ, while walking through the city.

The Location Visualiser

The Location Visualiser in LitLong is a web-based interface that provides a general overview of place name mentions in the final literary works we included in our dataset through the text mining and editing processes described in the previous section.

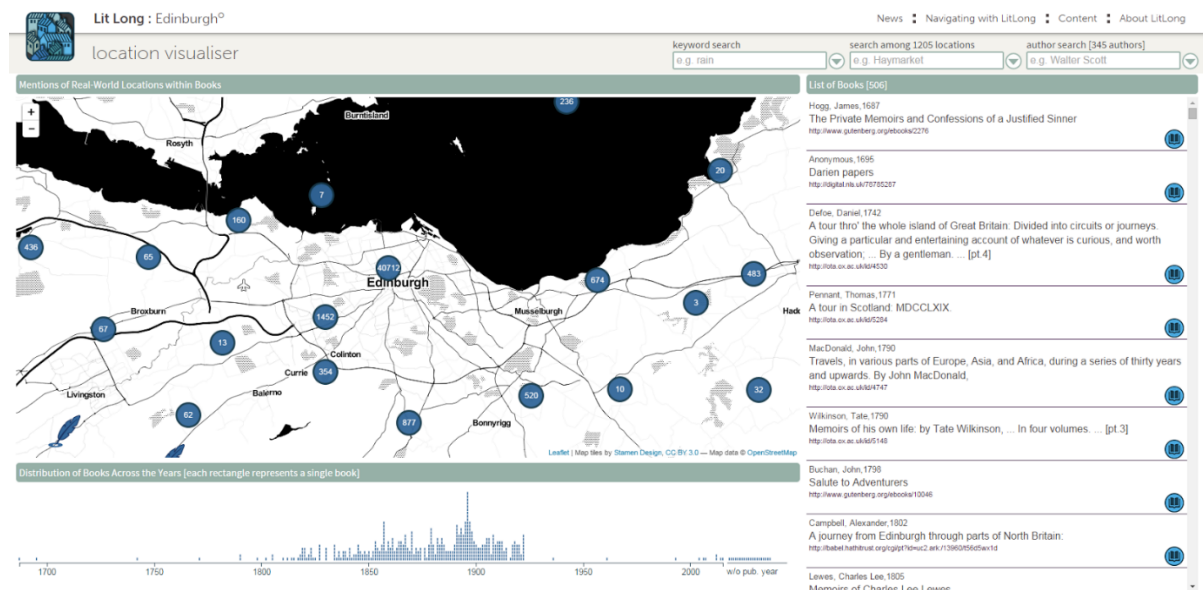


Fig.1: The Location Visualiser interface.

The interface consists of three main components: a map showing the real-word place name mentions included in the literary works, a timeline, showing the distribution of literature over time, and a book list, listing the literary works, their corresponding author, publication year, and a link to the digitized version of the book. All three views are interlinked: interacting with one automatically updates the others. Various search filters are provided to enable more targeted explorations by keyword, place name, or author. The Location Visualiser is implemented in JavaScript and PhD using standard libraries such as D3.js (for the general visualisation components) and Leaflet.js (for the map view). In the following we describe the different entry points the Location Visualiser facilitates into the Palimpsest literary collection and dataset.

One can start to explore data in the Location Visualiser with the map view. Here, place name mentions are presented on a conventional geospatial map based on their latitude and longitude.

Place mentions in close proximity are shown as clusters to avoid clutter in frequently mentioned areas. Individual location mentions are shown as a quill symbol and location mention clusters are shown as filled blue circles with a number representing the amount of unique place names they represent. For example, figure 2 shows four clusters of location mentions. Hovering over a cluster shows the geospatial area that it encompasses and a word cloud listing all corresponding place names with font sizes indicating their frequency of mention.

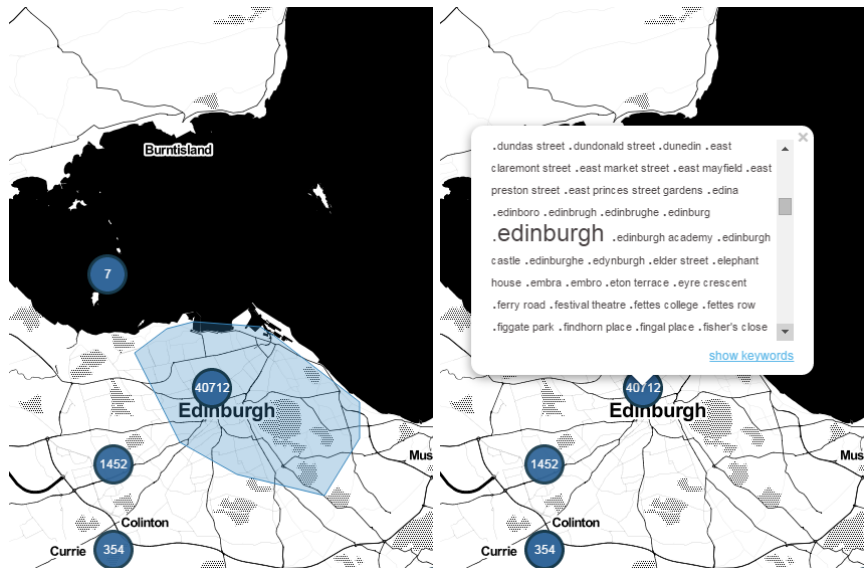


Fig 2. Location mention clusters in the Location Visualiser.



Fig.3

On demand, the word cloud also provides a glimpse of the literary context that the cluster encompasses. This begins to convey something of the context in which locations in this particular area were mentioned (see figure 3).

The inspection of location clusters from this high-level point of view provides an idea of the literary topology of Edinburgh in terms of the popularity of certain areas in and around the city within English literature. Place names referring to locations in the inner area of the city of Edinburgh are quite frequent, whereas places outside the inner city area are mentioned only in a few selected literary works (visible in the cluster of 7 mentions). Clusters with fewer than 20 location mentions will directly display the corresponding literary snippets, instead of showing an overview of place names or keywords. Users can browse these snippets one-by-one; for each snippet the book title, author, and publication year is provided and identified place names are highlighted within the snippet (see figure 4).



Fig. 4: Small clusters allow for browsing through corresponding snippets directly.

Zooming into the map enables a more fine-grained view of literary traces in and around the city. Clusters are broken down and details of the literary topology become visible (see figure 5). The higher the zoom level, the smaller the geospatial areas and range of place names that location clusters encompass (see figure 6).

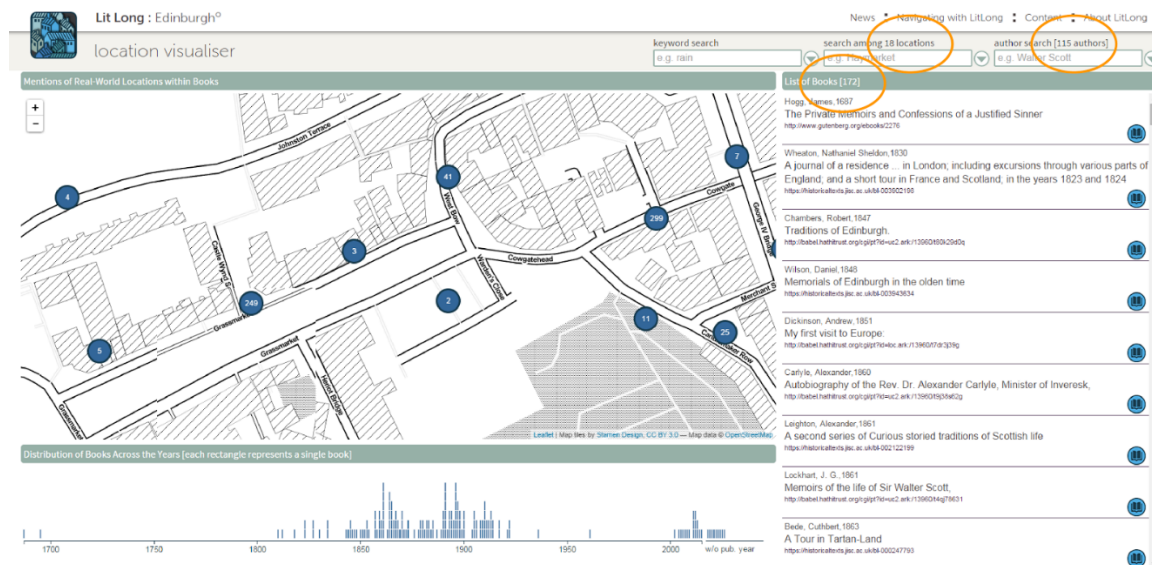


Fig. 5: Zooming into the map provides details about literary mentions of Edinburgh's Royal Mile area.

Interacting with the map (i.e., zooming and panning) updates the other views as well as the potential filter options. The map acts as a location filter and the numbers of locations and authors are directly adjusted accordingly. Similarly, the booklist and timeline view are updated to show only titles that include place names included in the currently visible map section.

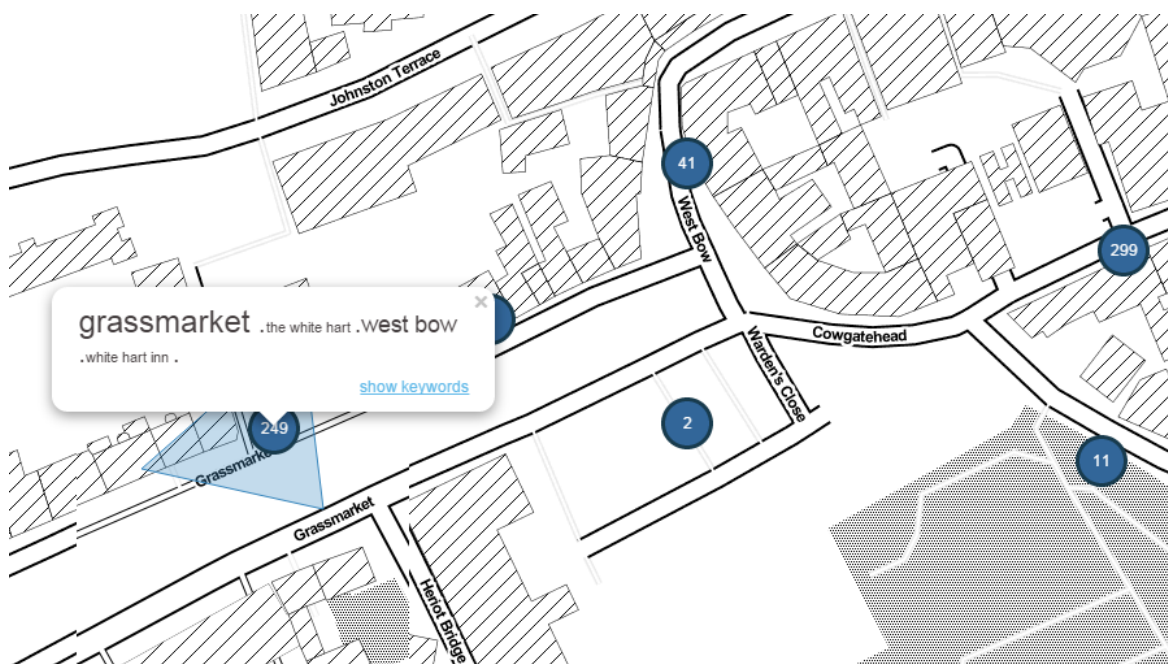


Fig. 6: Location clusters become smaller as the zoom level increases.

The dataset can also be explored on a book-by-book basis. The book list view simply lists each book included in our collection, including the paragraphs extracted from the book which contain Edinburgh place name mentions always corresponding to the area shown in the map as described above and (see figure 7).

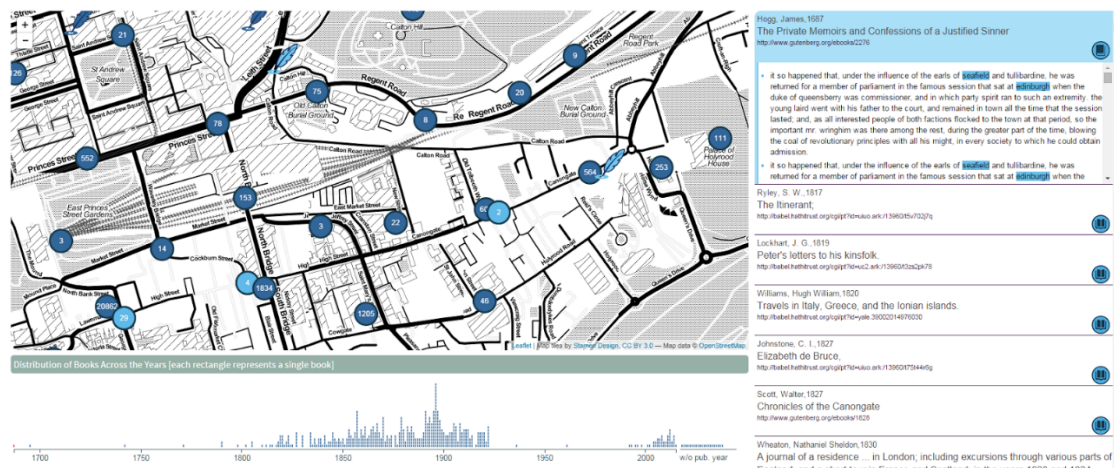


Fig. 7: Book list view (left) with the selected book's location mentions highlighted in light blue.

Hovering over a book item in the booklist highlights its corresponding place name mentions in the map (see figure 7). Again, multiple location mentions in close proximity are clustered, depending on the map's zoom-level. Selecting a book from the list via mouse click adds its location mentions as a permanent layer to the map. This can facilitate the comparison of selected literary works and their place name mentions (see figure 7).

The open-ended exploration approaches facilitated by the map and book list view are enhanced by a filtering mechanism that enables a targeted search of the collection by keyword, place name, and/or author. For example, entering an author of interest updates all views to only include books by this particular author. Autocomplete mechanisms are in place to provide direct-feedback on the availability of this particular author name in our dataset. Alternatively, a list of authors can be browsed to find authors of interest (see figure 8, left). Here, font sizes indicate the number of Edinburgh-related snippets included in the authors' works. Equivalent selection

mechanisms are available for the location and keyword search. The dropdown menus not only facilitate targeted search but can also convey an idea of the range of location names and authors that are available as well as the types of keywords that are included in the snippets of selected authors or locations. This augments the more specific exploration of particular location clusters in the map view, by providing an idea of keywords, location names, and authors corresponding to the entire map section currently in view.

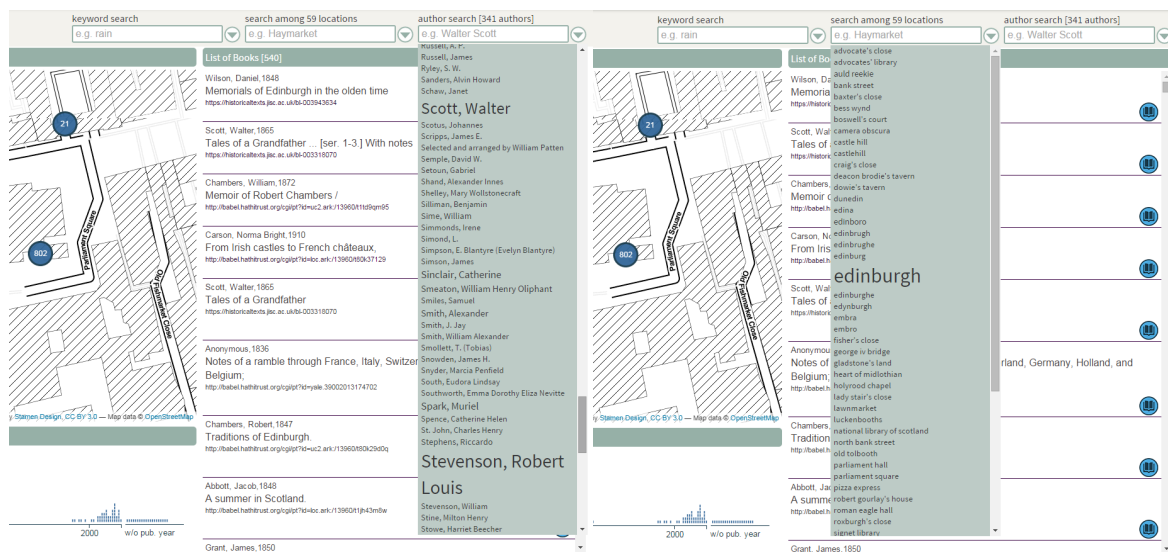


Fig. 8: Drop-down menus provide a glimpse of available authors, location mentions and contextual keywords.

The iOS App

The LitLong Mobile App (suitable for Apple iPhones and iPads) is the mobile counterpart of the Location Visualiser, providing opportunities to browse literary extracts while moving through the city of Edinburgh. It uses built-in location services of a mobile phone (GPS), to trigger location-specific requests to the LitLong database. The app was created in XCode, Apple's development environment. This enabled use of software development kits (SDKs) such as a map toolkit. The map toolkit allowed incorporation of Apple maps into the app and placing pins in the map at specified latitude and longitudes (lat-longs), which allows the user to feel comfortable and familiar with the LitLong app environment.

As it is not assumed that the user will always have connectivity to the internet, the data required in the app is stored inside the app itself rather than relying on a database accessed via the internet. This created both technical opportunities and challenges. Storing the data in the app enables quick local searching of data from nearby locations without the need to wait for the request to be sent to the server and then receive the required information back. However, as iOS apps require any data inside an app to be loaded only when the app is open, and the dataset used is quite extensive, the loading of the data had to be streamlined. The initial load data was reduced to only the data required for app setup (such as lat-longs, location names and number of extracts at each location), enabling the pin locations of nearby books to be dropped on the map in less than one second from initial launch. Then when more information from that location is required, such as the text extracts from books and their associated metadata, the data can then be loaded into the app on request. To split the data up in this way json files with the data for each location were stored and loaded on request.

App operation is simple, and deliberately intuitive. On tapping the app launch icon a splash screen is shown (see figure 9a). The user is then presented with a map and text boxes which show book information. The map can be used in two modes: either the user's current location serves to organise and select information, or the user taps on a specific book location shown with pen nibs on the map (see figure 9b for a screenshot of the interface). The selected information shows the location at the very top, then the title, author, year and a link directly to the original full text (if available) which if tapped opens in a web browser. The extract itself appears in a scrollable text field. If there is more than one extract at that location then the user can go through these by using the blue left and right arrows by the text.

If a pen-nib is tapped then information about that location is shown in a pop-up (see figure 10a), with the name of the location and the number of extracts at that location. Then if the user taps the information (i) button the distance from the user to that location is shown (see figure 10b). The map can be scrolled through to a different area by tapping and dragging, and zoomed in and out of

by tapping with two fingers and pinching. The user can also snap either to their current location by pressing the relevant button in the bottom left of the screen, or to the centre of Edinburgh by pressing the button in the bottom right. To further improve the browsing locations of interest, a maximized view of the map can be selected by going into the menu which can be accessed either by swiping in from the left or by pressing the “Menu” button on the top left (see figure 9c).

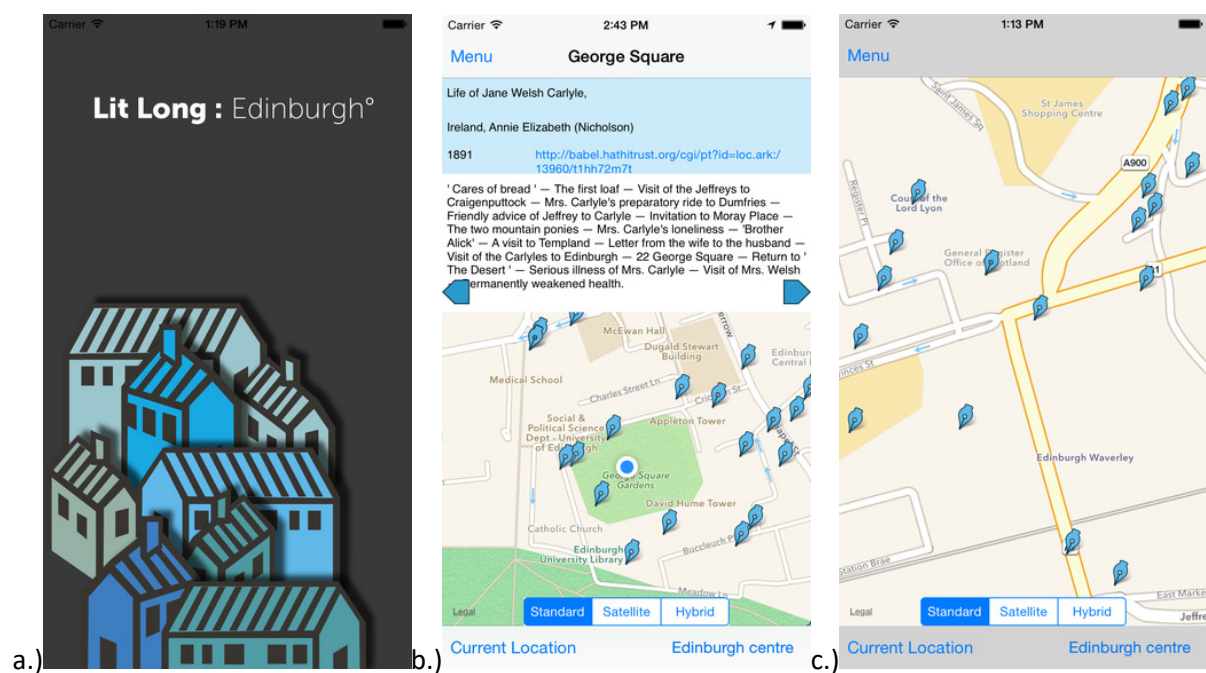


Fig. 9: Screenshots of the LitLong iOS App different views. a.) splash screen. b.) the main view, showing the map of the current location and current location. c.) the large map view.

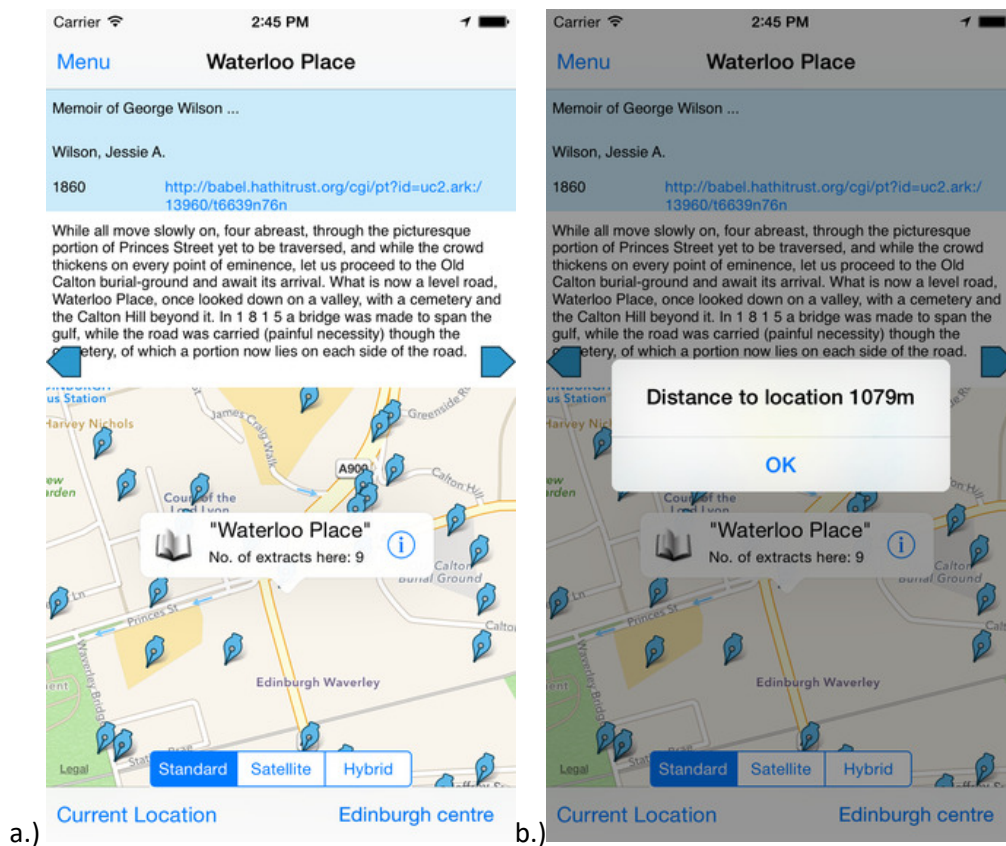


Fig. 10: Screenshots of the types of information available in the LitLong iOS App when selected a book location.

5: Unembarrassing the Eye

In his account of the presencing of place in literature, Stern Pultz Moslund, who has also been influenced by phenomenological and cognitive linguistic accounts, argues that topopoetic readings need to pay attention to the ways in which ‘the earth of the place juts through the cultural world of the novel’.²² This is not simply a one-way dynamic, as Moslund further describes: ‘Topopoetically, we approach the novel and its signifiers, not as standing *for* the place, ... but as standing *in* that very place ... The landscape in the work, or the place, seems to open up in front of us, behind us, above us, and beneath us’.²³

Reciprocal shaping occurs between the cultural and physical aspects of a place. Linguistic representations map on to underlying physical features and processes in ways that function

practically within a society and that are based on characteristics in ourselves and in the world. Yet even 'factual' accounts of places are necessarily coloured by sociocultural and cognitive biases. A fictional text's use of real world place-names makes the narrative a form of counterfactual scenario; an 'as if' account, that emerges through an author's (and their readers') drawing on and adding to the previous usages and associations of the place and its name. This means that there is always both continuity and difference between the accounts, which creates persisting and shifting patterns whose own continuities and divergences provide meaningful insights and potential inspiration.

Our initial explorations were driven by a general curiosity around the ways in which this key geocritical dimension of literary works might be made visible and explorable through digital means and methods. The creation of the database, visualisations and app have shown that it is possible to build what a number of scholars have termed a 'macroscope'. Tim Hitchcock has recently given this term a new twist, stressing not just the enlightening force of the large scale but equally the illuminating 'power of the particular'.²⁴ Such a machine combines high-level and large-scale explorations of literary works with the capability to attend to particular works, authors or locations. Our visualisations have focused on allowing users to pick their own path through the data, and to pursue alignments along a number of axes, such as place name, work, author, or keyword. These interactions involve either actual or virtual placement within a map of Edinburgh, but they don't easily amount to a totalising view of the terrain in its literary significance. Our maps, in other words, do not provide a single plane within which all of the data is visibly arrayed.

There are costs involved in this approach. The collections of digitised texts on which we have drawn are not comprehensive, despite their scale; at the same time, the difficulty that machine reading experiences in tracking inference or weak implicatures limits the extent to which the variety of ways in which place names can be contextualised or can be effectively mapped. This means that, to some extent, the large-scale approach has not made quantitative or statistical analysis any easier - as it is currently configured, LitLong does not readily permit users to gain a definitive, distantly read overview of spatial or temporal patterns of association and usage. Instead, the use of geolocation

and its expression in cartographic form makes it possible to disassemble texts into smaller units, our snippets, which can then be reassembled into new combinations on the basis of their geographical identity or contiguity. In its current form, LitLong offers a 'generous interface' - to use Mitchell Whitelaw's term (2012) - which furnishes its users with a rich experience of the particulars of literary place name usage, rather than a means of distant reading which meets the aspirations for that kind of analysis expressed most forcefully by Moretti.²⁵ As a macroscope, it privileges the singular encounter with the extract. It thus makes possible a range of new geocritical perspectives on the literature of single city, but it doesn't seek to establish a single or overall view. Indeed, it is as open to use by writers, readers and visitors as it is to academics - perhaps even more so.

Although our capacity to interpret literary place is not yet extended by these new media in the kind of panoramic or macroscopic way envisioned by some theoretical digital humanities' texts, the mining and visualisation of digitised texts allows both for new perspectives and for a continuation of the active reading practices which already inform scholarly and creative writing. Using LitLong it is possible to see, rather more immediately than otherwise, across the range and spread of place name usage associated with different authors – Walter Scott and Irvine Welsh, for example. The geolocated extracts also illuminate the diverse range of ways in which place names can operate across literary texts: as a grammatical entity of discourse; as a functional role in the story structure; as a means of achieving an aesthetic effect; as a representation of a thematic element or of a semantic complex.²⁶ The juxtaposition of the geolocated fragments of texts, like quotes in critical arguments or literary allusions, whilst not equivalent to nor a replacement for in-depth analysis of whole texts, can itself enable creative and scholarly insights.

In addition, our data could readily underpin further visualisations that might open up other lines of geocritical enquiry. For example, how do particular real-world place names mentioned in literary works relate to each other, beyond their geospatial attributes? Are certain locations frequently mentioned in relation to each other? How do the linear structure and temporal aspects of a narrative relate to the concept of place and space? And, in terms of context, how are particular

place mentions represented in terms of surrounding contextual and grammatical structures? How can we visualise qualitative and topological differences across particular place name mentions (e.g., the mention of a city, neighbourhood, road, or building)? New visualisation approaches are needed to address such questions, some of which may branch off from the concept of geospatial maps and towards more abstract visualisations where the place names are represented, for example, according to their appearance and proximity to other location mentions within the narrative. Similarly, contextual relations between place and, for example, emotion could be opened to exploration via alternative visualisation techniques.²⁷

Looking down on the city of Edinburgh from Calton Hill in his *Picturesque Notes*, Robert Louis Stevenson gives voice to the anxiety about informational superabundance that Westphal suggests will always defeat the attempt at geocritical analysis of culturally overdetermined sites. 'It is the character of such a prospect,' he says, 'to be full of change and things moving. The multiplicity embarrasses the eye; and the mind, among so much, suffers itself to grow absorbed with single points'.²⁸ Perhaps, with the advent of interfaces and forms of working in which close and distant reading might be usefully combined, it is no longer the case that the rich multiplicity of a site's literary geography need embarrass either the writer's or critic's eye.

¹ M. M. Bakhtin and Michael Holquist, *The Dialogic Imagination: Four Essays*, Austin: University of Texas Press, 1981.

² Michel de Certeau and Steven Rendall, *The Practice of Everyday Life*, Berkeley: University of California Press, 1984.

³ Yi-fu Tuan, *Space And Place: the Perspective of Experience*, Minneapolis: University of Minnesota Press, 1977.

⁴ Henri Lefebvre, *The Production of Space*, Oxford: Blackwell, 1991.

⁵ Robert Tally, 'Textual Geographies: The Real-and-Imagined Spaces of Literature, Criticism, and Theory'. *Reconstruction* 14, 2014

⁶ Ibid, 2.

⁷ Bertrand Westphal and Robert Tally, *Geocriticism: Real and Fictional Spaces*, New York: Palgrave Macmillan, 2011; Bertrand Westphal, *The Plausible World a Geocritical Approach to Space, Place, and Maps*, Basingstoke: Palgrave Macmillan, 2013.

⁸ Franco Moretti, *Atlas Of the European Novel, 1800-1900*, London: Verso, 1998; *Distant Reading*, London: Verso, 2013; *Graphs, Maps, Trees: Abstract Models for a Literary History*, London: Verso, 2005.

⁹ See 'Mapping the Lakes: A Literary GIS', <http://www.lancaster.ac.uk/mappingthelakes/>; 'A Literary Atlas of Europe', <http://www.literaturatlas.eu/en/>; Shelley Fisher Fishkin, "'Deep Maps": A Brief for Digital Palimpsest Mapping Projects (DPMPs, or "Deep Maps")', *Journal of Transnational American Studies* 3, 2011; 'The Modernist Versions Project', <http://web.uvic.ca/~mvp1922/>.

¹⁰ Robert Tally, 'Foreword', in Westphal, *Geocriticism*, xii

¹¹ Tally, 'Textual Geographies', 2.

¹² Laurence Hutton, *Literary Landmarks of Edinburgh*, New York: Harper & Brothers, 1891, xiii.

¹³ Ibid.

¹⁴ James Loxley, Beatrice Alex, Miranda Anderson, David Harris-Birtill, Claire Grover, Uta Hinrichs, Jon Oberlander, Lisa Otty, Aaron Quigley, James Reid, and Tara Thomson. *Lit Long: Edinburgh*, <http://litlong.org/>, 2015.

¹⁵ Tally, 'Foreword', in Westphal, *Geocriticism*, xii.

¹⁶ T. Strohman, D. Metzler, H. Turtle, and W.B. Croft, *Indri: A language-model based search engine for complex queries* (extended version), CIIR Technical Report, 2005.

¹⁷ Beatrice Alex, Claire Grover, Jon Oberlander, Ke Zhou and Uta Hinrichs, 'Palimpsest: Improving assisted curation of loco-specific literature,' in *Proceedings of DH2015*, Sydney, 2015.

¹⁸ Ted Underwood, 'Understanding Genre in a Collection of a Million Volumes', Interim Report, December 2014, <http://dx.doi.org/10.6084/m9.figshare.1281251>.

¹⁹ See Beatrice Alex, et al, 'Adapting The Edinburgh Geoparser for Historical Georeferencing,' *International Journal of Humanities and Arts Computing Intl J Humanities & Arts Computing* 9, 2015, 15–35.

²⁰ D. Ganguly, J. Leveling, and G. J.F. Jones, 'Automatic prediction of text aesthetics and interestingness', in *25th International Conference on Computational Linguistics (COLING 2014)*, 23-29.

²¹ Irvine Welsh, *Trainspotting*, London: Jonathan Cape, 1993, 228.

²² Sten Pultz Moslund, 'The Presencing of Place in Literature: Toward an Embodied Topopoetic Mode of Reading,' in Robert Tally, ed., *Geocritical Explorations: Space, Place and Mapping in Literary and Cultural Studies*, NY, Palgrave Macmillan: 2011, 29-43, at 38.

²³ Ibid, 39.

²⁴ Tim Hitchcock, 'Big Data, Small Data and Meaning.' *Historyonics*, 9 November 2014; http://historyonics.blogspot.co.uk/2014/11/big-data-small-data-and-meaning_9.html.

²⁵ Mitchell Whitelaw, 'Towards Generous Interfaces for Archival Collections,' *Comma* 2012, 123–132.

²⁶ Miranda Anderson and James Loxley, 'The Poetics of Place-Names in Digitised Literary Edinburgh,' in David Cooper, Christopher Donaldson, and Patricia Murrieta-Flores, eds, *Literary Mapping in the Digital Age*, London: Routledge, 2016.

²⁷ See, for example, the essays in Christian Nold, *Emotional Cartography: Technologies of the Self*, S.l.: s.n., 2009.

²⁸ Robert Louis Stevenson, *Edinburgh: Picturesque Notes*, London: Seeley, Jackson, and Halliday, 1879, 31.