




Article

Making BERT Feel at Home

Modelling Domestic Space in 19th-Century British and Irish Fiction

Svenja Guhr¹ 
Jessica Monaco² 
Alexander Sherman³ 
Matt Warner² 
Mark Algee-Hewitt² 

1. iSchool, University of California , Berkeley, USA.
2. Literary Lab, Stanford University , Palo Alto, USA.
3. English Department, UT Austin , Austin, USA.

Citation

Svenja Guhr, Jessica Monaco, Alexander Sherman, Matt Warner, and Mark Algee-Hewitt (2025). "Making BERT Feel at Home. Modelling Domestic Space in 19th-Century British and Irish Fiction". In: *Journal of Computational Literary Studies* 4 (1). 10.48694/jcls.4164

Date published 2025-09-30

Date accepted 2025-09-11

Date received 2025-02-07

Keywords

British and Irish fiction, 19th Century, Victorian, domesticity, space, classification

License

CC BY 4.0 

Reviewers

Anouk Lang, Sandeep Soni,
Anonymous Reviewer

Note

This paper has been submitted to the conference track of JCLS. It has been peer reviewed and accepted for presentation and discussion at the 4th Annual Conference of Computational Literary Studies at Krakow, Poland, in July 2025.

Abstract. We introduce a novel approach to detecting domestic space in literary texts beyond explicit spatial markers like "home" or "house." Using a pre-trained English BERT model fine-tuned on manually annotated passages from a corpus of 19th-century British and Irish novels, we develop a method to operationalize and quantify domesticity in fiction. Our model captures the nuances of domestic space by analyzing contextual and relational cues rather than relying solely on toponymic and other explicit references. This approach offers new insights into the representation of space in literature, revealing the fluid and dynamic nature of domesticity in 19th-century British and Irish fiction.

1. Introduction

She went upstairs, emerging all at once into the full morning sunshine in the hall, which dazzled and appalled her. [...] She went into Clara's room first. [...] Clara's maid was seated, fast asleep, before a table on which a candle was burning pitifully in the full daylight. The room looked trim and still as a room does which has not been occupied in that early brightness. The maid woke with a shiver as Mrs. Burton entered. "Oh, Miss Clara, I beg your pardon," she said. "It is no matter. My daughter will not want you tonight. Go to bed, Jane," said Mrs. Burton. (*At His Gates*, Margaret Oliphant)

What makes space domestic in fiction? Is it the mention of keywords like "home" or "room"? Is it the presence of characters discussing private matters? Or is space domestic when characters are engaged in private or intimate interactions, as in *At His Gates* when Mrs. Burton checks on her daughter's room while talking to the housemaid? Space occupies an important place in literary theory, and domestic space in particular gains importance in 19th-century fiction. As, for example, Davidoff and Hall (1987) detail, the Victorian period is marked by "separate spheres," in which men participated in public and professional life while women were responsible for the home as the central organizers of domestic life, combining both physical space and domestic ideology. In fiction, as Cohen (2017) explains, portrayals of domesticity both criticized and upheld domestic

ideologies as novels populated their homes with characters ranging from the angelic Agnes Wickfield in Dickens' *David Copperfield* to the villainous figures of sensation fiction. Such attention has resulted in a significant body of criticism on domesticity in Victorian fiction, including but not limited to Armstrong (1990)'s additional focus on class in *Desire and Domestic Fiction*, Freedgood (2006)'s emphasis on empire and materiality in *The Ideas in Things*, and Marcus (2007)'s work on friendship and sexuality in *Between Women*. However, these studies and others tend to prioritize addressing the concept of domesticity over a strict account of domestic space. In both fiction and literary criticism, spatial information offers a concrete link between domestic settings and ideologies and allows readers to orient themselves as characters move through the fictional worlds they inhabit. Such settings are also implicated in themes of gender, class, and colonialism. Our project, therefore, sought to operationalize space in fiction (especially domestic space) in order to trace the patterns of domesticity and its associated cultural constructs through the British and Irish 19th-century novel.

The operationalization of space has a long history in the context of computational literary studies. Moretti (1999) and Piatti (2016) concentrate on the importance of geographic plotting in the construction of narrative meaning, while Ryan et al. (2016) and Wilkens (2013) have applied computational methods to map fictional settings onto real-world entities. Other examples include Bamman et al. (2019), who annotated and automated the recognition of named spatial entities in BookNLP (Bamman 2021), as well as Bologna (2020) and Schumacher (2023), who similarly operationalize space by identifying Bamman et al. (2019)'s sets of spatial keyword classes (e.g., GPE, LOC, FAC, etc.) using machine learning techniques. These approaches rely on explicit spatial references, such as named entities like toponyms or spatial entities such as "marketplace" or "sitting-room," which are the focus of the most recent work by Kababgi et al. (2024), who fine-tune a BERT language model to automatically detect and recognize non-named spatial entities (NNSEs) from manually annotated training data. Using sentence-based annotations, they first identify sentences containing NNSEs and then classify them as "rural," "urban," "natural," or "interior." While these methods have proven adept at detecting explicitly spatialized passages, passages without these entities are often difficult to spatially identify. This problem underscores the challenge of identifying implicit space.

In this paper, we introduce a new method for the automated detection of both explicit and implicit domestic space in English-language fiction based on the probability of a passage being set in domestic space. Our approach offers a departure from the implicit ideological or ontological framework of previous approaches – where domestic space is predefined as a static concept – by adopting a phenomenological one. Instead of asking "Is this space domestic?", we switch to the question "How likely is it that the passage is set in domestic space?" This change allows us to explore how domesticity manifests in ways that challenge traditional assumptions as we identify the domestic qualities of unexpected or liminal spaces like gardens, carriages, or even ships. To that end, we propose the calculation of a "domesticity score": a score based on the probability assigned to a passage by a fine-tuned English BERT classifier (trained on manually annotated data) of a passage being set in "domestic space." This modeling approach offers new possibilities for the analysis of fictional spaces that are not explicitly described but are discursively constructed through dialogue, context, and emotional tone.

In our paper, we first describe our operationalization of “domestic space” and then detail the annotation process that we used to operationalize our corpus of 19th-century British and Irish fiction. Second, we introduce a multilingual transformer model fine-tuned to compute the probability of a passage being set in “domestic” space through a two-step classification task performed on six-sentence passages. Using our model, we calculate the “domesticity score” for each passage in our corpus, which we can then summarize across each novel. We then provide an analysis of chosen texts by canonical authors to offer a new perspective on implicit domestic space. This intervention opens new opportunities for analyzing space, character, and plot in fiction.

2. Operationalizing Domestic Space

Our project to identify domestic space in 19th-century British and Irish fiction began with a derivative approach to annotation-based concept operationalization recommended by Pichler and Reiter (2022). We similarly did not start from a specific working definition of explicitly and implicitly represented “domestic space” in fiction. Rather, we approached the concept through approximation, using exploratory annotation, inter-annotator agreement calculation, and discussions, resulting in the iterative development of a decision tree for the annotation task. Our rationale was that, while theoretical frameworks in narratology operationalize space via narrated action involving characters or descriptions of physical environments, textual clues to setting are often absent from narrative discourse (see Fludernik and Keen 2014; Ryan 2014). For instance, the spatiality of an event may be inherited from descriptions in previous scenes (frequent in novels with long dialogues), remain implicit in character interactions, or be altogether absent in reflective passages that are narrated non-spatially. Focusing on examples of domestic space, we recognized that domestic spatiality is not clearly bound to entities, but is a fluid literary concept that varies contextually. For instance, a garden may sometimes function as a domestic space within a narrative about children playing or adults discussing romantic entanglements during a stroll, becoming an extension of the private sphere of the home. In other contexts, however, gardens can be part of publicly accessible parks, whether or not they are adjacent to the homes of the wealthy. This example only emphasizes the difficulty of setting the boundaries for a clear definition of domestic space in fiction in a way that captures its full ideological, historical, and cultural dimensions.

In contrast to existing definitions that risk excluding the ambiguities that make domestic space so central to fiction, we adopted an inductive approach to operationalizing domestic space. Rather than imposing a fixed definition, we fine-tuned a language model on agreed-upon examples of domestic space, allowing the model to infer patterns and associations that characterize settings. By approaching the classification of space with machine learning methods based on contextual embeddings, we offer a fluid definition of domestic space through a “domesticity score” that measures the probability of a passage being set in domestic space against it being set in another type of space or being non-spatial. In our manual annotation process, we included passages set in living rooms, kitchens, bedrooms, etc. that provide strong indicators of what constitutes a domestic space beyond named or non-named entities, e.g., through resolved coreferences, deictics, or other explicit and implicit clues detectable by human annotators. In the same way, we also included passages with explicit settings that are not domestic (e.g., battles, ships at

sea, or carriage rides). By using these clear examples as part of a training sample, we enabled the model to detect domesticity even in passages that lack overt spatial markers. In this way, we could extrapolate from explicit examples of domestic space to implicit examples recognized by the model as sharing all of the same features except the explicit references to domestic space. The model's ability to generalize from training data allows it to classify all of the passages in our corpus and reveal patterns of domesticity.

We aimed to categorize passages into two primary classes: "domestic space" and "other." The choice to limit the classification to these two classes was driven by the nature of our future research interest: By focusing on domesticity, we aimed to isolate passages of interest for broader inquiries into themes of gender, colonialism, and social hierarchies in Victorian fiction. Attempts to differentiate the class "other" into subcategories (e.g., public spaces, natural landscapes, or non-spatial passage) proved impractical for several reasons. On the one hand, annotators often struggled to achieve consensus on subcategories, given the inherent fluidity and overlapping boundaries of non-domestic spaces, as well as the limited context given in the annotation passages. For example, from a six-sentence passage, it was often impossible to tell if the passage was spatialized non-domestically or just non-spatialized. On the other hand, our primary goal was accurately recognizing domestic spaces, not exhaustively classifying different types of spaces.

To transform the abstract concept of domestic space into measurable units, we defined these units as fixed-length six-sentence text segments. This segmentation allowed us to systematically apply annotations and later model predictions across the corpus. We relied on intersubjective interpretation during the annotation process. This system involved iteratively creating a set of guidelines that balanced theoretical rigor with practical applicability. Annotators were tasked with identifying passages that unambiguously depicted either domestic settings, such as interiors of homes, or non-domestic settings such as workplaces or otherwise public settings. Ambiguous or marginal cases were excluded. This operationalization ensured that the training data for our model represented the clearest possible examples of both domesticity and non-domesticity to minimize uncertainty in the machine learning process.

3. Data and Method

3.1 Data Preprocessing

We used a corpus of 19th-century British and Irish novels (see [Table 1](#)), sourced from the University of Illinois libraries and Chadwyck-Healey Nineteenth-Century Fiction collection (Chadwyck-Healey Literature Collections and ProQuest 2016). The corpus represents a curated selection of literary texts, including canonical and lesser-known literary prose. Additionally, the collection offers detailed metadata, such as publication dates and author information, which enables diachronic and comparative analyses. Although not the largest corpus available and not strictly representative of 19th-century

Texts	2,865
Words in total	557,097,804
Individual authors (+ 126 “unknown”)	1,250
British authors	1,226
Irish authors	24
Texts by Irish authors	118
Time period	1748–1899
Six-sentence passages	3,684,727
Manually annotated passages	1,227
“domestic space” passages	521
“other” passages	678
“trash” passages	28

Table 1: Research corpus metadata summary. 126 texts have “unknown” authors.

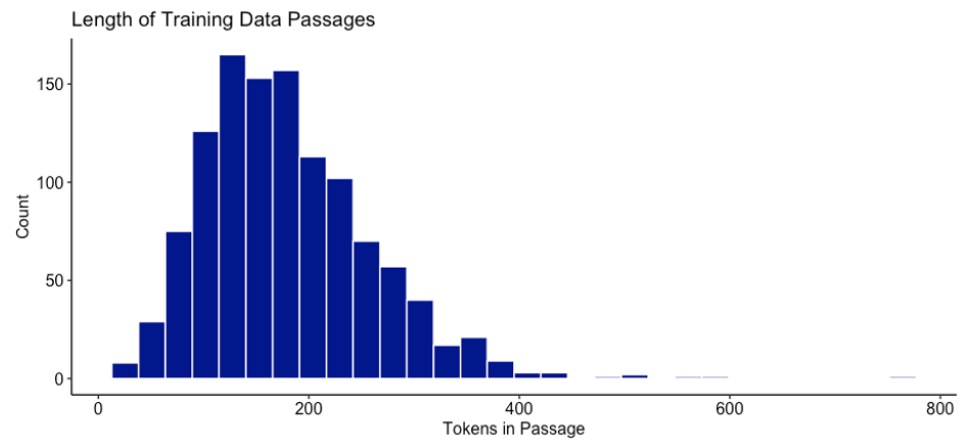


Figure 1: Varying passage lengths in the training corpus. Passage length did not correlate with the class choice of the model.

novelistic prose, our corpus offers relatively clean OCR¹ (many of the texts were hand-keyed) and a sample of both canonical and non-canonical texts.

One of our first decisions for the project lied in our chosen resolution for the passages we wanted to classify. Chapters made up of multiple scenes would be too long to classify as domestic or not (the action of a chapter might move from a bedroom to a garden to a carriage), while sentences would be too short (within a given chapter, only a few sentences actually contain information on the setting). Paragraphs, although closest to our desired resolution, are too inconsistent in length (particularly when representing dialogue) for reliable classification with our transformer model. In our previous close-reading approaches, six-sentence passages proved to be the Goldilocks zone: long enough to get enough spatial information, short enough to be mostly one space and to be read and classified quickly enough by human readers while manually tagging the passages (see Figure 1). Furthermore, the six sentences strike a balance between granularity and context. They capture enough of the narrative to identify domestic space without introducing excessive noise. During the annotation process, six-sentence passages provided sufficient context for human annotators to make informed decisions about spatial settings that aligned with the model’s training needs.

1. The sample passages quoted in this paper are taken directly from the digital versions of the literary texts. To illustrate the classification results transparently and without concealing the OCR quality, the passages are reproduced exactly as they appear in the corpus, including OCR errors and additional white spaces.

3.2 Manual Annotation

Following the recommended workflow for annotation guideline creation by Reiter (2020), we defined annotation classes and developed a decision tree (see Figure 2) giving annotators an ordered set of decisions to follow before declaring a passage to be set in “domestic space” or “other.” As a third class for manual annotation, we defined “trash” for passages that either contained paratextual material (such as bibliographic references or advertisements) or were unreadable to human annotators due to excessive OCR errors or foreign language². The annotation guidelines were iteratively developed through their application, annotation evaluation, discussion, and guideline refinement to ensure clarity, consistency, and alignment with the conceptual framework. We define the classes as follows:

- **“domestic space”**: passages set in clear, unambiguous domestic settings, such as interiors of homes,
- **“other”**: passages set in non-domestic or ambiguous spaces, including public places, natural landscapes, or spaces where the setting was unclear or non-spatial as in reflective passages or summaries,
- **“trash”**: passages with poor OCR quality, foreign language, or extra-textual or paratextual elements.

Five experts trained in literary studies and two student assistants manually annotated 1,375 passages. The passages were selected partly because they contained a domestic seed term, such as “kitchen”³, and partly at random from the corpus. Each passage was annotated by at least two independent annotators, resulting in a total of 3,657 individual annotations. Following the decision tree for manual annotation, each passage was tagged with one of the three classes: “domestic,” “other,” or “trash.” In an earlier version of the annotation guidelines, we added the category “I don’t know” (“IDK”) to the decision tree to distinguish between passages that were unambiguously non-domestic or non-spatial from passages that gave no information on their spatiality at all.

As part of the iterative development of the annotation decision tree, we added a majority decision step when encountering “IDK” passages containing information on more than one space or non-spatial elements mixed with some spatial information. Namely, for passages containing more than one space, annotators were told to classify the passage based on the location of the majority of its sentences. For instance, in the case of a six-sentence passage in example 1, the first four sentences cover a setting in domestic space as the characters prepare to go outside. Their exit is narrated at the end of the passage in the last two sentences as they walk “towards the gardens.” The agreed-upon decision for this passage by all annotators was the class “domestic space.” At a later stage, we incorporated “IDK” into the *ex negativo* class “other” to focus our annotation on the detection of domestic space. Example 2 contains a passage that has been manually annotated as “trash.” Out of all 1,375 passages annotated, about 30% of the passages

2. By labeling foreign language passages as “trash,” we identified foreign language novels in the corpus that did not meet the selection requirement of being English-language texts. Although our chosen multilingual transformer model accurately identified “domestic space” and “other” passages independently of language, we excluded these texts from future analyses focusing solely on English-language prose.

3. A list of the used annotated seed words can be found in the GitHub repository.

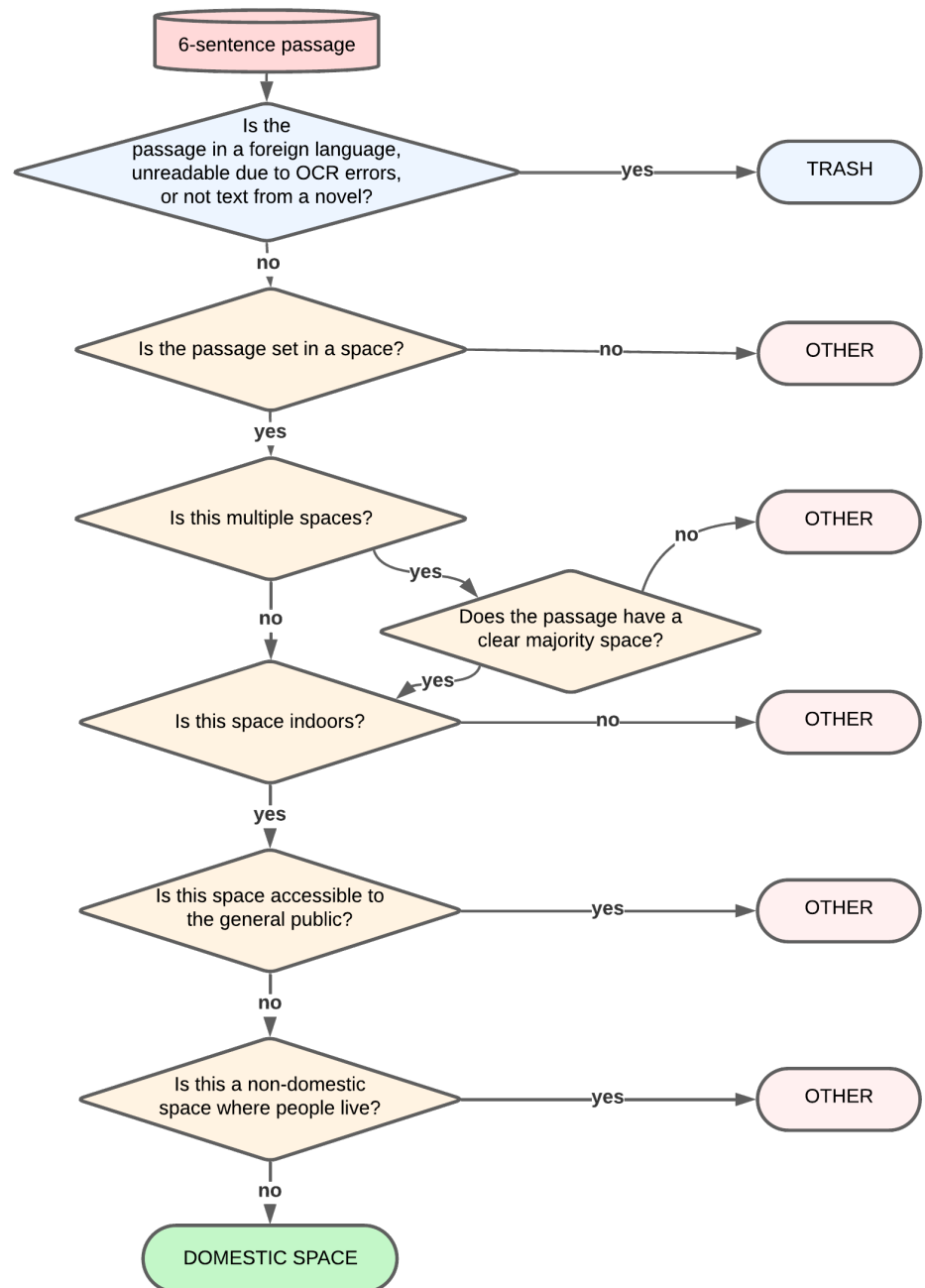


Figure 2: Decision tree for the manual annotation of a passage as “domestic space,” “other,” or “trash.”

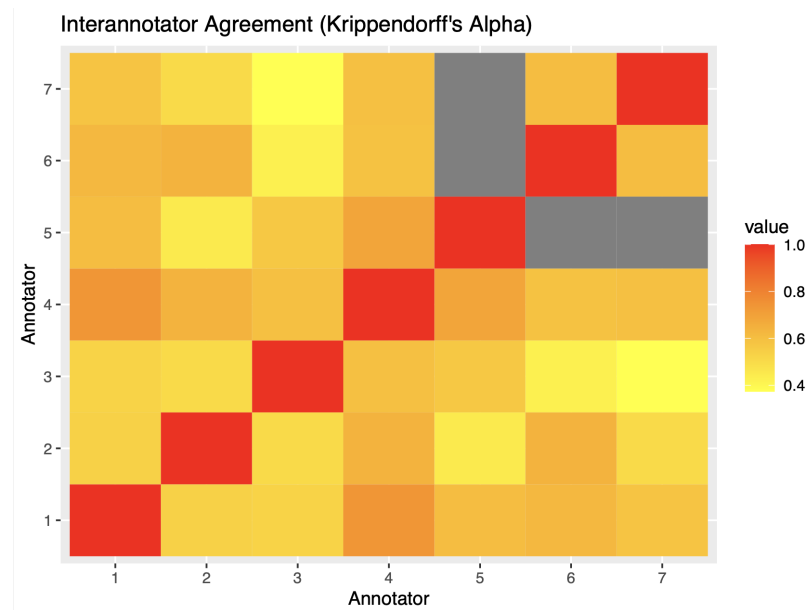


Figure 3: Heat map showing the (dis)agreement between annotators calculated with Krippendorff's Alpha.

were classified as “domestic space,” 67% as “other,” and 3% as “trash” giving us an initial benchmark against which to measure the automated performance of the model.

Ex. 1 It is warm and mild now, and we shall be back in time for luncheon, I will just get my hat.” He went into his bedroom as he spoke, and after a moment came back with his hat in his hand. John had left the room and was standing just outside the door. As Sir Lionel came through the sitting-room, he watched him furtively, but closely; and as soon as he was fairly in the corridor, John shut the door, and, forgetting his usual deference, led the way briskly through the porch. They walked towards the gardens; but presently John said: “I fear you will have some further trouble with James, I hope he will go this afternoon.” “I hope so, these scenes of howling and supplicating are very tiresome.”

(passage from *Riding out the Gale* by Annette Lyster labeled as “domestic”)

Ex. 2 THE LAWS OF WAR AFFECTING COMMERCE AND SHIPPING. By H. BYERLEY THOMSON, of the Inner Temple. Second Edition, greatly enlarged. 8vo, price 45. %d. boards. LECTURES ON the ENGLISH HUMOURISTS OF THE 18th CENTURY. By W. M.

(metatextual element labeled as “trash”)

3.3 Validation of the Annotations through Ground Truth

To assess the reliability of our manual annotations, we calculated the inter-annotator agreement (IAA) using Krippendorff's Alpha, a statistical measure for categorical data annotated by more than two annotators (Krippendorff 2018). The overall Krippendorff's Alpha for our annotations was 0.58 across five annotators, which is below the standard threshold of 0.8, but consistent with the inherent subjectivity and ambiguity observed in similar literary annotation tasks (see Figure 3).

Despite the relatively low Alpha, the qualitative comparison of the manual annotations did not reveal any systematic deviations or rogue annotators. Instead, disagreement

was evenly distributed, reflecting the underlying complexity of identifying fictional space. Assessing annotation quality, however, extends beyond inter-annotator agreement. While Baledent et al. (2022) question whether high agreement necessarily ensures accuracy, a key challenge remains: Annotators may converge on errors, making intersubjective consensus yield a lower quality annotation than ground truth. To evaluate the validity of our annotated data, we established a set of ground truth annotations for passages where the consensus converged on one annotation rather than the other, despite the absence of explicit spatial markers in the extracted text segment.

As Pichler and Reiter (2022, 14) explain, validity serves as the critical “link between theory and measurement,” allowing researchers to evaluate whether their methods genuinely align with their conceptual objective. Similarly, Krippendorff (2018, 361) emphasizes that a measurement instrument is valid “if it measures what it claims it measures.”

In literary studies, intersubjectively recognized annotations – those agreed upon by multiple annotators – are considered a robust measure of validity. As gold annotations, they are used as the basis for text analysis and interpretation, as well as for training models for automation. However, during our annotation process, we observed a key limitation: While high inter-annotator agreement confirmed the reliability of our classifications, the annotations themselves did not always capture the true spatial context of a passage. On the contrary, given a six-sentence passage without an explicit lexical marker for spatial information, the annotators had to decide whether the passage was set in domestic space based on the given information, such as private dialogues or intimate actions, which are more likely to be set in domestic space than in public space and have to be spatial by default since characters are present. Nevertheless, in the discussion rounds, the annotators often could not justify their annotation decision by referring to elements on the textual surface, even when an intersubjective annotation decision was given. Consider the example 3, where the annotators initially only saw six sentences of the dialogue, which they agreed contained little spatial information and suggested an “other” classification. The passages in set brackets (presented here in abbreviated form), however, show the surroundings of the dialogue, taken from the novel, which clarify that it actually takes place in a domestic space.

Ex. 3 {But he had not had time to finish his sentence before the door of the house was thrown open, and Stephanie Harcourt appeared upon the threshold.

“Bella” she cried to her friend hysterically, “it is all over. I am dismissed without salary, and I can’t even pay you my share of the week’s rent ! The sooner I go to the Tombs with that scoundrel the better!”

“Hush, hush, dear ! there is a stranger present,” said Miss Vavasour compassionately. [...] “My poor child, how came you to marry him?”

“I can’t tell you that. I was frightened into it in a way that you would hardly understand. Only, thank heaven, I am now delivered from him.”}

“But after his two years’ incarceration are over, he will come out again and claim you.”

“I will have broken the chain by that time. I will have gone far away where he shall never find me.”

“And you met Cortes in San Francisco?”

“Yes, sir.”

“And that scoundrel Sandie Macpherson had some hand in your marrying him?”
 { The girl’s cheek became as white as ashes. “Who has told you that?”
 “No one. I guessed it”}
 (*Phyllida. A Life Drama* by Florence Lean, 1882)

This is the key difference between gold annotations and ground truth. While we achieved a high inter-annotator agreement in manually classifying six-sentence passages as “domestic” or “other,” we wondered whether our intersubjective class choices actually represented valid annotation choices for the passages when we took the greater context of the passage into account (context that was unavailable to our annotators and which would be unavailable to our model). Accordingly, we decided to go beyond our gold annotations and manually verify the spatial setting of a given passage by looking at where each passage fit within the novel, and by searching outside of the passage (before or after) for contextual information about the actual space in which the passage is set.

We conducted this contextual validation on a subset of 15 passages, with additional annotations informed by the surrounding text. This process revealed some new findings: Many passages that were initially labeled as “other” in the gold annotations were reclassified as “domestic space.” For instance, dialogues that appeared spatially ambiguous within the passage, e.g., due to the lack of any spatial marker in the dialogue itself, were often revealed to occur in domestic settings when viewed in context. Going back and forth several pages before and after the passage (sometimes up to 30 pages needed for long dialogue passages and on average ten minutes needed for the classification of one passage⁴) allowed us to find spatial referents for our target passages, and thus enabled a ground truth classification for the six-sentence passage. Passages containing dialogues or transitional scenes (e.g., characters moving between spaces) were the most likely to be reclassified. These results highlight the challenges of detecting implicit domestic space based on limited textual context alone and underscore the importance of ground truth annotations for classification tasks beyond the gold annotations that annotators agree on. While gold annotations provide a standardized and efficient means of generating training data, ground truth annotations offer more fidelity to the actual text being annotated.

However, creating ground truth annotations is even more expensive than creating gold annotations because of the extra labor involved in tracking down the contextualizing information. Furthermore, for the purpose of automating the classification task, we had to consider that the state-of-the-art transformer models we use are also constrained to a limited context. Therefore, the decision to use six-sentence passages proved to be an appropriate heuristic: While the passages are short enough for manual examination, they provide a relatively high level of contextual information for the classification task. Since we could not provide a large number of ground truth classifications for training, we kept the ground truth annotations out of the training set.

4. In comparison, the preparation of gold annotations took approximately 30 seconds for reading and deciding a class for a six-sentence passage.

4. Automation: Make BERT Feel at Home

Transformer-based architectures have emerged as a preferred approach for classification tasks in computational literary studies (CLS), offering greater transparency than large language models (LLMs), which are often optimized for language generation rather than classification (see e.g. Bamman et al. (2024)). Pre-trained models from the BERT family (Devlin et al. 2019) have been successfully applied in various literary and linguistic classification tasks, including genre attribution (van Zundert et al. 2022), character gender identification (Schumacher et al. 2022), emotion classification in plays (Dennerlein et al. 2023), and the detection of dubitative passages (Parigini and Kestemont 2022). For automated space recognition, recent studies have demonstrated the superior performance of fine-tuned BERT-based models over LLMs such as GPT-3.5 and GPT-4 (Kababgi et al. 2024; Soni et al. 2023). Given these findings, we selected a BERT-family model for our sequence classification task, specifically the TensorFlow Universal Sentence Encoder (USE) model (Yang et al. 2021).

As we describe above, unlike prior work on spatial classification that relies on entity detection (Kababgi et al. 2024; Soni et al. 2023), our study shifts the focus from explicit spatial markers to the implicit discursive construction of domestic spaces. To implement our approach, we fine-tuned a pre-trained English BERT model from TensorFlow Hub on our manually annotated training data. Initially, we used TensorFlow’s BERT_en_uncased preprocessor with an English BERT model pre-trained on Wikipedia and BooksCorpus and fine-tuned on the Multi-Genre Natural Language Inference (MNLI) dataset (Devlin et al. 2019; Google 2023a). While BERT_en_uncased is widely used for NLP tasks and designed for token-level tasks like question answering and named entity recognition, capturing bidirectional word context, the Universal Sentence Encoder (USE) generates fixed-size sentence embeddings, making it more effective for semantic similarity and sentence classification. Consequently, the USE model offers superior performance in complex, higher-order tasks (such as classifying space). It is also multilingual, offering an additional advantage for passages containing foreign language words (a semi-regular occurrence in 19th-century novels) and outperformed the BERT model for our classification task. For these reasons, we ultimately selected USE due to its strong performance in sentence-level embeddings and its effectiveness in transfer learning, particularly in low-data settings. The model employs a Transformer-based sentence encoding architecture that computes context-aware representations of words while preserving both word order and surrounding context (Cer et al. 2018; Google 2023b). This enables effective sentence-level transfer learning for our six-sentence segments, providing higher classification performance with a small set of training data.

To develop a classifier for detecting domestic space in British and Irish fiction, we fine-tuned an up-to-date (2023) USE model using TensorFlow and Keras. The training process followed a two-step classification approach. First, we trained a binary classifier to filter out “trash” passages with the understanding that these would not be relevant for further classification. This first model was trained on manually labeled data, where passages were categorized as either “trash” or “not trash.” The data was preprocessed using the USE multilingual preprocessor, tokenized, and passed through the USE encoder. The model was trained with categorical cross-entropy loss and optimized

using the Adam optimizer⁵, incorporating early stopping to prevent overfitting. Once trained, this model was used to filter out irrelevant passages from the dataset, ensuring that only meaningful textual segments were passed to the second classification step.

The second model classified the remaining passages into “domestic space” or “other” categories. This model was trained in a similar manner, using a labeled dataset where passages were tagged accordingly. Again, we used the USE preprocessor and encoder to generate sentence-level embeddings, which were then fed into a neural network with a dropout layer to mitigate overfitting. The trained model was saved for reuse, allowing for batch classification of unseen textual data.

4.1 Prediction

After training, the models were deployed to classify new texts. To facilitate prediction on unseen data, we employed a sequential two-model pipeline as dual binary classifications allowed for the development of separate specialized models for recognizing trash and identifying domestic passages, respectively. This enabled us to ensure high-quality predictions while leveraging the strengths of USE’s sentence-level embeddings for transfer learning in a low-data setting. The input consists of an Excel or CSV file containing segmented literary plain texts, where each cell contains a six-sentence passage generated through prior segmentation using the spaCy sentence splitter (Montani et al. 2023; SpaCy 2024).

The first stage of the prediction pipeline is trash detection, where a trash detection model assigns a probability score indicating whether a passage is classified as “trash” or “not trash.” Segments with a high probability of being “not trash” are retained for further analysis. The filtered output consists solely of text segments deemed relevant for domesticity classification. The remaining passages were then analyzed by the domestic space classifier, which assigned probabilities to each passage being “domestic space” or “other.” In this second stage, we predicted the domesticity score for each six-sentence passage in the corpus using a rolling-window approach. The model reads the cleaned dataframe and predicts a domesticity score for each six-sentence segment, determining the likelihood of its setting being domestic. The prediction operates independently for each segment, meaning that the surrounding textual context – both preceding and following passages – is not considered. The classification is based exclusively on the content within each individual cell. This two-step approach proved more successful than a three-way classification task (“domestic space” vs. “other” vs. “trash”). The classification results were compiled into structured tables for further analysis.

4.2 Evaluation of the Model Performance

Model evaluation was conducted using a held-out test set alongside ground-truthed annotations. To avoid sampling bias introduced by initial keyword-based selection, the held-out test set was randomly sampled from the full corpus, ensuring a more representative and independent evaluation. Our assessment of results – at both the novel and passage levels – suggests alignment with established critical expectations.

5. Adam is an algorithm that combines the advantages of the Adaptive Gradient Algorithm (AdaGrad) and Root Mean Square Propagation (RMSProp) by adjusting learning rates for each parameter based on estimates of first and second moments of the gradients (Kingma and Ba 2015).

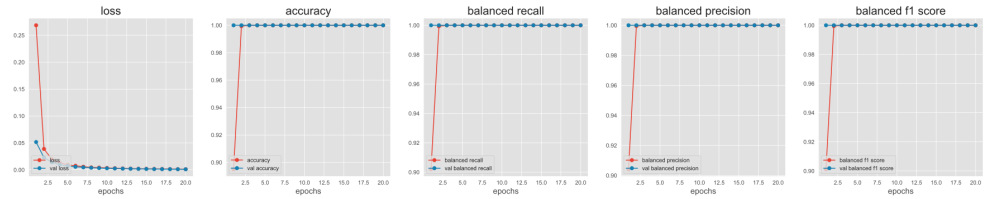


Figure 4: Performance of the “Trash Detector” model.

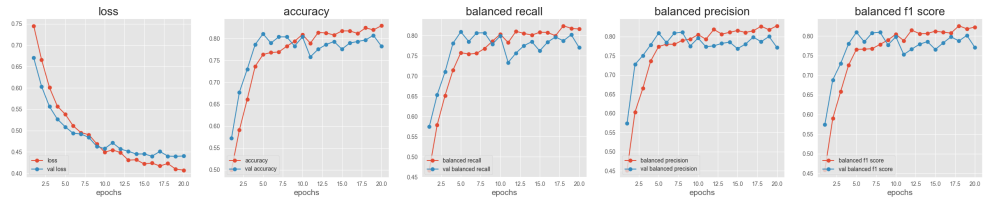


Figure 5: Performance of the “Domestic Space Detector” model.

Training performance was visualized over multiple epochs to monitor improvements (see [Figure 4](#) and [Figure 5](#)), and early stopping was applied to optimize performance. The model was evaluated using key metrics, including categorical accuracy, recall, precision, and F1-score (see [Table 2](#)), followed by a sample-based error analysis.

Trash Detector. In the first step of our pipeline, the most frequent misclassification of trash occurs when the trash detector fails to filter out foreign-language passages (French in particular) that were manually labeled as “trash” in the test data. This indicates that the model was not explicitly trained to use language as a distinguishing criterion for “trash.” For example, passages with non-English dialogue, but also segments of foreign language texts (see [example 5](#)), which were missed during the manual cleaning of the data set, remain for the second prediction step. However, with the transition from BERT uncased, pre-trained on English texts, to the multilingual Universal Sentence Encoder, the model retains the ability to predict whether a passage is set in a domestic setting. Accordingly, this ‘misclassification’ proves to be a desirable feature of the model, despite its disagreement with our annotation guidelines. Another common misclassification occurs with segments of low OCR quality, which manual annotators labeled as “trash,” but the trash detector did not predict as being too bad to be excluded (see [example 6](#)).

Consequently, given that foreign-language passages are relatively rare in the dataset (but still present despite manual checks for foreign-language texts), and that the model’s ability to accurately classify passages of low OCR quality based on their setting is an advantage rather than a disadvantage; this limitation does not affect the overall

	Trash Detector	Domestic Space Prediction
Accuracy	1.00	0.8159
Recall	1.00	0.8105
Precision	1.00	0.8092
F1-score	1.00	0.8097
Loss	0.0012	0.4288

Table 2: Model evaluation results for the “Trash Detector” and the domestic space prediction.

effectiveness of the pipeline. On the contrary, the trash detector still performs well on these segments, outperforming⁶ human annotators in these cases.

Ex. 4 FALSE STEPS 1 64 XIII. WANT OF MONEY 179 XIV. IN THE GLOAMING 1 97 CHATTER PAGE XV. [...]

(True positive: index manually labeled as “trash,” automatically predicted as “trash” with a probability of 0.15 (“not trash”) to 0.85 (“trash”))

Ex. 5 Enfin, ils se sont tous ruinée, et un M. Stanley a acheté le bien. Si je ne me trompé, il était le premier mari de Lady Clarancourt et il lui a laissé le Manoir, mais seulement en usufruit. [...]

(False negative: manually labeled as “trash,” automatically predicted as rather “not trash” with a probability of 0.57 (“not trash”) to 0.43 (“trash”))

Ex. 6 [...] He addressed a most affectionate letter to tubert, informing him of the death of Mrs. Sedley, and the total change which had ^ken, place; adding, that in consequence towliieb be added, lfe->fird\$ ^^uite ^i-^6M- ^sAedi; and oa b& irectrm ^shi-^u}d^(-^lb plearore yfeld bis wife up t6^4A^^^^fie tberr added, that d-^a the i ^fidnlfiicf^lt^lSft ^wc^M not allow her to Wril^bifB^^^^- ^dlf, she liad requested him^U> ^petfonsli tIMIt office for her. %He conctaded” by de^iUfg his sdn to address all lett^^in^ililu^

(False negative: manually labeled as “trash,” automatically predicted as rather “not trash” with a probability of 0.54 (“not trash”) to 0.46 (“trash”))

Prediction of Domesticity. We validated the domesticity prediction model by selecting an additional random sample of 120 passages from the corpus, manually annotating them, and comparing the results with the model’s classifications. The model and annotators aligned in 71% of cases (85 out of 120), surpassing the initial inter-annotator agreement (IAA). Further analysis of the model’s probability scores reinforces these findings. In 84 instances, the model assigned a high-confidence probability (either above 70% or below 30%) for a passage being categorized as “domestic,” with annotators agreeing 82% of the time. For passages where the model showed greater uncertainty (probabilities between 40% and 60%), agreement dropped to 44%. These results indicate that most discrepancies arose in passages the model itself recognized as ambiguous. In a further validation step, we did an error analysis of the predicted domesticity scores of the segments that were included as part of the ground truth data set (see [subsection 3.3](#)).

The predicted domesticity scores for passages labeled as “domestic” in the ground truth data reveal intriguing patterns: Among the 19 passages identified as pure dialogue without explicit spatial markers, the model assigned an average domesticity score of 0.45 with a standard deviation of 0.2. Notably, 15 of these 19 passages received a score below 60%, suggesting that the model frequently registered uncertainty when encountering dialogue without explicit spatial cues. Conversely, the seven passages categorized as pure dialogue in “other” settings showed the model’s tendency to correctly assign them to non-domestic spaces. These passages had an average domesticity score of 0.26,

6. While we do not suggest that the model outperforms human annotators in theory-driven classification tasks, in the specific case of the “trash” category, characterized primarily by textual noise rather than interpretive ambiguity, the model shows greater consistency, particularly in detecting low-quality OCR passages that annotators often disagreed on.

corresponding to a 0.74 probability of being “other,” with a standard deviation of 0.18. Moreover, five of the seven passages received a low domesticity score (<40% “domestic”, >60% “other”), indicating a clearer classification.

These findings raise interesting questions about the role of dialogue in spatial classification. While dialogue alone does not strongly signal domesticity, it appears that the model struggles more with assigning high domesticity scores to dialogue-heavy segments without explicit spatial markers. This suggests that contextual cues beyond six-sentence windows, such as speaker identity, dialogue patterns, or adjacent descriptions, may play a critical role in determining domesticity. Further investigation of dialogue structure as a latent feature in domesticity classification will be discussed in [subsection 5.3](#).

Finally, we acknowledge that the model is highly overspecialized to detect 19th-century domesticity, as it has been trained specifically for this purpose. For example, if applied to texts from Latin American Boom fiction in translation, it would still attempt to assign domesticity scores using the criteria it has learned from 19th-century novels despite contextual differences. However, this historical specificity aligns with the goals of our project, which aims to capture and analyze domesticity as it was conceptualized in 19th-century British and Irish fiction.

4.3 Domesticity Score

Since the output of our classification tasks consists of numerical values between 0 and 1, the received numbers provide a way to identify passages with a high probability of being set in “domestic space” or “other” (or of being “trash” for the first classification task respectively) and can be taken directly as a score indicating the relative domesticity of the passage. With this approach, we are able to provide information about the likelihood of a passage being set in “domestic space” or “other” rather than providing forced binary decisions for one class. As a result, passages of ambiguous spatial nature are present (and identified as such), as well as passages that tend toward one of the two classes. Based on this, each passage considered in the second classification task was assigned a domesticity score between 0 and 1. The output of the classification task is a dataframe in which each classified passage is identified by a distinguishing passage ID and the classification value for being set in “domestic space” or “other,” enriched with metadata about the title of the text from which the passage was taken, the author’s name, and the publication date.

The analysis of domesticity scores highlights key patterns in how the model interprets domestic space in fiction. Passages with the highest domesticity scores, such as those from *The Ill-tempered Cousin* by Frances Elliot (see example 7) and *Ombra* by Margaret Oliphant (see example 8), exhibit rich domestic imagery, explicit spatial markers, and detailed descriptions of household activities. For example, in *The Ill-tempered Cousin*, the passage’s focus on household disorder, personal belongings, and family interactions contributed to its nearly perfect domesticity score of 0.996. Similarly, the passage from *Ombra*, with a score of 0.978, features a cozy, well-defined domestic setting, emphasizing warmth, comfort, and familial intimacy. In contrast, passages with low domesticity scores often lacked clear spatial markers or were dominated by dialogue without explicit references to domestic settings. The model showed greater uncertainty when processing such ambiguous segments, particularly in cases where dialogue occurred without

contextual grounding. This suggests that while the model effectively identifies overtly domestic scenes, it – like many readers – struggles with less explicitly defined spaces, reinforcing the need for further analysis of latent features such as dialogue patterns and indirect spatial cues.

Ex. 7 Everything in the house that morning was in confusion. The housemaid had put coarse sheets on Lady Danvers’ bed, and forgotten the muslin curtains to the window. [...] A letter, too, had come from John Bauer (how many hours the excellent John had spent over its composition in the solitude of Wood’s Green, who can say?) telling of the deep impression Miss Escott had made on him, and requesting his aunt’s permission to return, “Only to be allowed to look at her,” wrote honest John, in a strictly business hand, with dots on all the i’s, and the t’s crossed to such a nicety, it would have been a pleasure to look at them, to anyone less worried than Aunt Amelia. [...] (Passage from *The Ill-tempered Cousin* by Frances Elliot, automatically predicted as “domestic” with a probability of 0.996)

Ex. 8 Mrs. Anderson’s room was a large one; opening into that of Ombra on the one side, and into an ante-room, which they could sit in, or dress in, or read and write in, for it was furnished for all uses. It was a petit appartement, charmingly shut in and cosy, one of the best set of rooms in the house, which Kate had specially chosen for her aunt. Here the mother and daughter met one night after a very tranquil day, over the fire in the central room. [...] Ombra came in from her own room in her dressing-gown with her dusky hair over her shoulders. Dusky were her looks altogether, like evening in a Winter’s twilight. (Passage from *Ombra* by Margaret Oliphant, automatically predicted as “domestic” with a probability of 0.978)

5. Analysis and Results

In this section, we compile the predicted domesticity scores across the texts in our corpus and visualize them diachronically to get a new perspective on domesticity within British and Irish literary history over time (see [subsection 5.1](#)). We then focus on authors (see [subsection 5.2](#)) and also address the challenges posed by dialogic passages (see [subsection 5.3](#)) to detect domestic spaces. As the proportion of “domestic” to “other” classifications in our automated classification echoes the percentages found by our annotators (described above), we take this as an additional validation for our model. The strong performance of our model in detecting the specific space class “domestic” based on manually labeled data highlights the potential of our classification approach and suggests that similar techniques could be successfully applied to other space classification tasks, such as identifying urban settings in detective fiction or automobile scenes in American short stories.

5.1 Domesticity and Literary History

Our domesticity score is useful for classifying passages as “domestic” or “other” when they are above and below our 70% and 30% thresholds, respectively. However, given the ambiguity of the scores between these cutoffs, the score itself is less meaningful for summarizing domesticity across a novel. Accordingly, we used the binary classifications

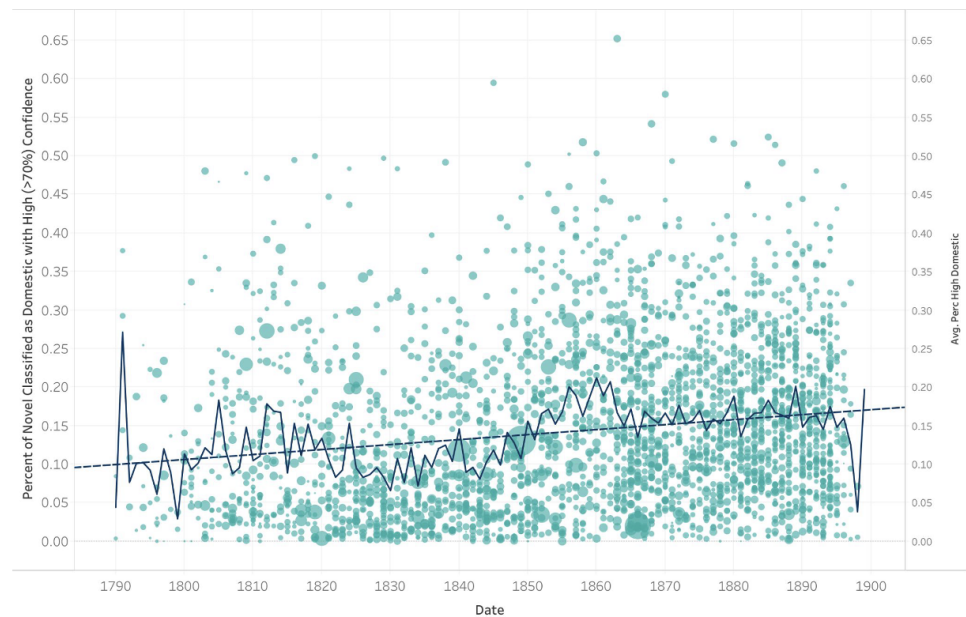


Figure 6: Domesticity score trendline of the 19th-century novel corpus. Due to the limited data points provided for the respective years, the beginning and end of the line plot are not representative.

based on the domestic score to calculate the percentage of passages each novel contained that were classified as “domestic” with a greater than 70% probability. We then visualized these percentages per novel to gain insights into the diachronic development of domesticity across the corpus. The novel with the highest percentage of passages predicted as “domestic” (highest dot in Figure 6 at 0.65) is Julia Kavanagh’s *Queen Mab* (1863) – an Irish author known for her “fashionably domestic [...] style” and writing for young women readers (Sutherland 1989, 343). The next highest point is at 0.60, which is British author Elizabeth Missing Sewell’s novel *Gertrude* (1845), primarily set in the home of the female protagonist and stressing the importance of familial responsibilities (Frerichs 1974). The third and fourth highest dots are again written by Julia Kavanagh, namely *Silvia* (1870) with 0.58 and *Dora* (1868) with 0.54.

The trendline provides a lens to examine the shifting prevalence of the domestic in different novelistic genres over time. For instance, the late 18th-century, characterized by slightly lower domesticity levels, coincides with the popularity of Gothic romances and travel narratives set abroad. In the 1810s, Jane Austen’s domestic novels emerge, followed by the rise of historical and Newgate fiction in the 1820s and 1830s. From 1850 to 1870, there is a noticeable increase in domesticity, likely linked to the prominence of domestic spaces in both realist novels and sensation fiction. Toward the end of the 19th-century, the growing popularity of adventure fiction, which by default does not represent domesticity, reshapes the Victorian novel, with the trendline reflecting this shift.

5.2 Domesticity and Canonicity

For the authors writing in the British Romantic period, from the last decades of the 18th-century through the earliest decades of the 19th, the points representing their

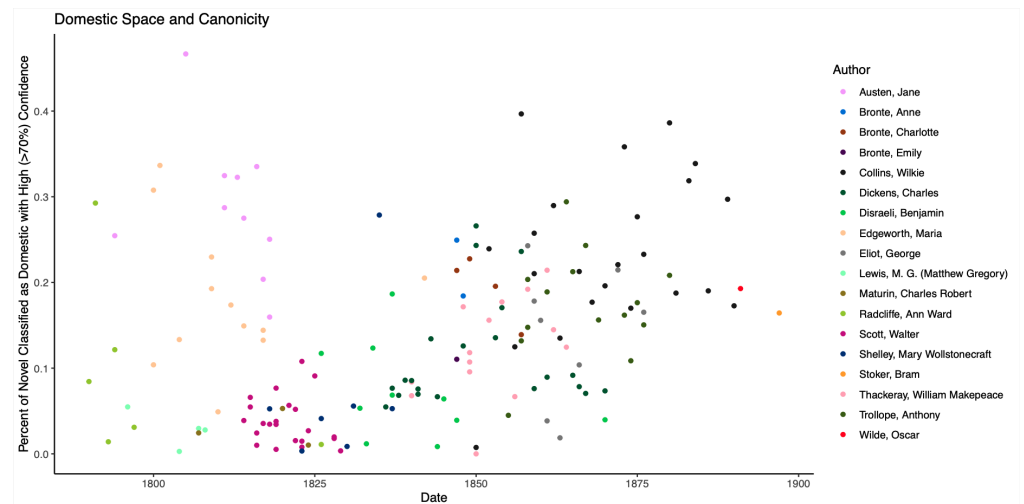


Figure 7: Visualization of a set of canonized authors’ texts and the percentage of passages with a likelihood of being domestic above 70%.

novels tend to form distinct clusters. These clusters also tend to correspond to particular novelistic genres. Ann Radcliffe and Matthew Lewis, whose points group together in the bottom left-hand corner, are both writers of Gothic fiction. Gothic novels in the Romantic period often take place in castles (which could be tagged as domestic spaces according to our annotation guidelines, referring to the public or private access to the room in question) or convents (which, despite that people live in them, were always tagged as non-domestic within our annotation guidelines). Ann Radcliffe, in particular, is known for her long, descriptive scenes of sublime landscapes and outdoor travel. To the right of their clusters, the points representing novels by Walter Scott also form a distinct group. Walter Scott’s historical novels tend to focus on public spaces and represent the characters’ experiences within large historical events (see also Lukács (1983)). A slight exception to this pattern of highly-clustered authors is Jane Austen, whose marriage plots spend so much time in houses that two of her novels – *Mansfield Park* (1814) at 0.28 and *Northanger Abbey* (1817) at 0.16, which is an old abbey converted into a domestic space – are named after them⁷. The location of the biggest outlier among her works, *The Watsons* (1805) at 0.46, seems to be, in part, a factor of length, since it was never published and exists only as novel fragment of 21,505 words.

In the Victorian period (1837–1901), realism and sensation fiction dominate the graph. Although their plotting differs – realism prioritizes everyday life, whereas sensation fiction foregrounds exceptional crimes and secrets – both genres often take place in homes. That being said, unlike the canonical authors represented in the earlier part of the century, authors like Charles Dickens, George Eliot, and Wilkie Collins are often spread out across a range of percentages for passages classified as highly domestic. For Dickens, for example, the most “highly domestic” novel is *David Copperfield* (1850) at 0.27, the one that, fittingly, has Angel-in-the-House Agnes Wickfield. However, most of Dickens’s novels hover around 0.05 to 0.15 and show investment in representing

7. While *Northanger Abbey* is indeed titled after a domestic site, much of the novel’s action actually unfolds in public and quasi-public settings like Bath, with the abbey serving more as a site of symbolic and imagined significance than as the primary narrative location. However, despite this detail, the Austen texts provide a very high number of domestic space passages on average in relation to the other authors’ texts, underscoring her sustained focus on the interior and private spheres.

both work and home environments. Even *Bleak House*, a novel named directly after two houses with that exact name and, arguably, after many other bleak homes represented alongside them, is only slightly more “highly domestic” at about 0.14 than the other Dickens novels represented by the points on either side of it. Given Dickens’ interest in representing the courts and the slums of London in *Bleak House*, this does not come as a surprise. Eliot’s novels hover mostly around 0.2, with some above and some below; *Middlemarch* (1872) at 0.21, known for being a canonical example of Victorian realism, includes several marriage plots and their respective domestic spaces, but it is also steeped in the politics and labor of the town of Middlemarch and the surrounding countryside. Of Collins’s sensation novels, *The Dead Secret* (1857) at 0.4, is the most “highly domestic” according to the model’s classifications; like many works of sensation fiction, this novel centers an inheritance plot and themes of family and illegitimacy.

The placement of some points on the visualization may be surprising. Oscar Wilde’s *The Picture of Dorian Gray* (1890) and Bram Stoker’s *Dracula* (1897) could be identified as about 0.19 and 0.16 “highly domestic,” respectively. Although early iterations of the Gothic novel as practiced by Radcliffe and Lewis rarely take place in domestic spaces, in the more urban Gothic of Wilde and Stoker, these Gothic plot lines more often do; take, for example, the location of Dorian’s portrait in his own home.

5.3 Domesticity and Dialogue

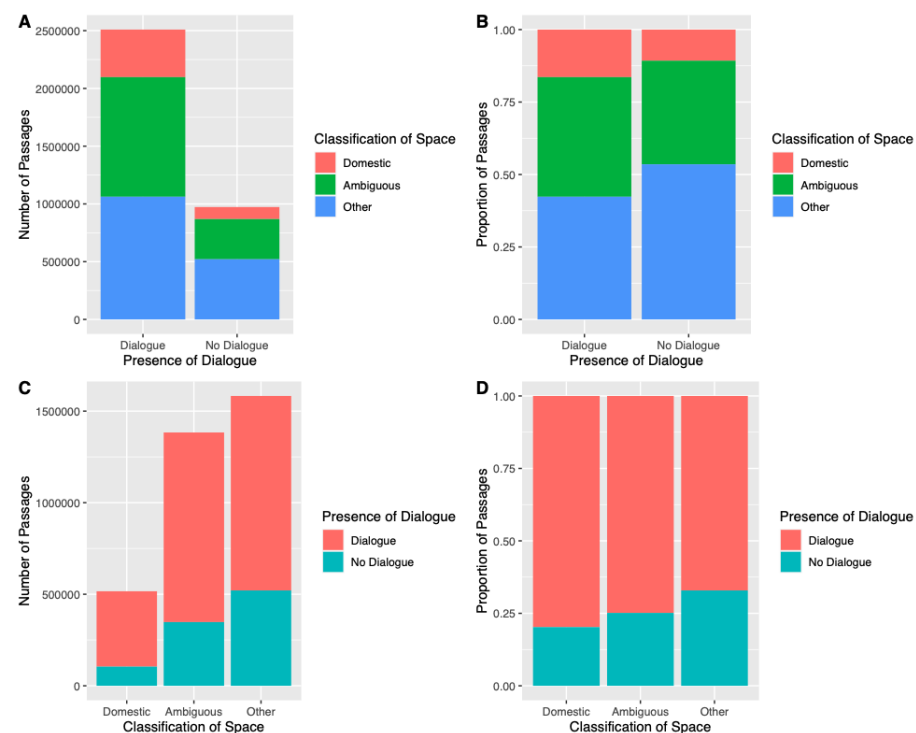


Figure 8: Visualization of the absolute numbers and proportions of dialogue in domestic space.

In our annotation process, we noticed that dialogue was a common source of difficulty. Passages containing dialogue often seemed to be set in domestic spaces, but they lacked any explicit signs of their location, and we thus often could not definitively tag them for inclusion in our training data when restricted to the six-sentence passages. Our

	Dialogue	No Dialogue	Totals
Domestic Space	411,718	104,462	516,180
Ambiguous	1,035,524	347,836	1,383,360
Other	1,062,109	520,714	1,582,823
Totals	2,509,351	973,012	3,482,363

Table 3: Number of passages in domestic space, as classified by the model with probability > 0.7, compared to the number of passages containing at least some dialogue, as estimated by the presence of quotation marks.

intuitions aligned with literary critical arguments about the correlation between household interiors and dialogue in domestic fiction. We were vindicated when, during our ground-truthing process (see [subsection 3.3](#)), we found passages consisting wholly of dialogue that our model correctly identified as set in domestic space (see [example 9](#)).

Ex. 9 “Doubtless, my dear,” said Casaubon, with a slight bow. “The notes I have here made will want sifting, and you can, if you please, extract them under my direction.” “And all your notes,” said Dorothea, whose heart had already burned within her on this subject, so that now she could not help speaking with her tongue. “All those rows of volumes – will you not now do what you used to speak of? – will you not make up your mind what part of them you will use, and begin to write the book which will make your vast knowledge useful to the world? I will write to your dictation, or I will copy and extract what you tell me: I can be of no other use.”

(Passage from *Middlemarch* by George Eliot labeled as “domestic” with a probability of 0.85)

To investigate this relationship between dialogue and domestic space further, we conducted a short exploratory study, where we found that passages containing dialogue were more likely to be set in domestic space and vice versa, a strong signal for their connection. As a proxy for the presence of dialogue, we found all the passages that contained single or double quotation marks, excluding those used as apostrophes. This method is somewhat imperfect: It misses passages from the middle of monologues, while catching those that might contain only a short portion dialogue at the end or beginning. It also encounters some problems due to OCR, dialogue without quotations, and quotation marks at the end of passages. In a sample of 100 passages, the method’s recall for finding passages with dialogue was 0.92, the precision was 0.9, and the F-score was 0.91. However, we judged these results sufficient for an exploratory study of the correlation.

Our results (see [Figure 8](#) and [Table 3](#)) show that dialogue is present, even predominant, across all spatial categories. However, passages with dialogue are 53% more likely to be in domestic spaces than those without dialogue, and passages in domestic space are 19% more likely to include dialogue compared to passages set in unambiguously non-domestic spaces. As we move from non-domestic to ambiguous and finally domestic spaces, there is more and more dialogue. The bidirectional relationship, contrast with non-domestic spaces, and high number of observations across categories imply a strong connection between domestic space and dialogue. Future work might explore the underlying factors in this relationship; we hypothesize, based on an analysis of the words distinctive of domestic spaces, that the prevalence of names, personal address, and family titles in dialogue plays a role. But our brief analysis here underlines our larger methodological arguments. Literary texts represent space much more complexly than

just through mentions of place names and spatial terms, including in dialogues between characters that do not include any explicit spatial information yet still signal a domestic setting. Our method is able to detect these pervasive, nuanced, and fundamental aspects of literary space.

6. Discussion: From Domestic Space to Domesticity

Our analysis highlights the complex interplay between domestic space and domesticity, emphasizing that expected domesticity does not always align with physical domestic spaces. While houses frequently serve as markers of domestic settings, domesticity is not solely confined to them. The model successfully identified high domesticity scores in traditionally domestic environments, yet it also revealed instances of unexpected domesticity in unconventional locations such as gardens, carriages, and even ships (for ships see the examples 10 and 11). These results are technically incorrect in identifying domestic spaces, yet also perceptively pinpoint both how spaces become marked as domestic and what it ultimately means for a space to be domestic. Domesticity extends beyond physical structures, emerging instead through relational and behavioral cues, such as familial interactions, caregiving, or moments of emotional intimacy.

Ex. 10 “My dear Merlin,” Power said to him, as he ascended over the ship’s side, “have you obliged me?” “I have, Power,” Merlin said. It was a little matter of no consequence, only of considerate kindness, but it was now that the easy terms of friendship commenced.

(Passage from *Tregarthen Hall* by James Garland labeled as “domestic” with a probability of 0.72)

Ex. 11 The ladies did not retire after dinner, but reclined on the one and only sofa in the cabin, whilst the gentlemen chatted over their wine, interrupted only by the warrant officer on duty coming and going with messages concerning the ship, when, and for a moment only, the wind was heard, and a very partial knowledge of the weather understood, so snug and comfortable was the cabin of the *Sylvia* and so agreeable the company. Tea was served with a continuous flow of conversation; anticipations of home meetings and happy days to come were the subjects. Merlin had now to leave his cabin to attend to his ship; the hour for weighing the anchor had arrived, and he examined his ship and gave forth his order to weigh. Very soon Mrs. Power and Iselen were made sensible of the delusion they had been under that the *Sylvia’s* cabin was a quiet place, and a sea voyage could not be so terrible a thing after all Merlin stepped below and told them he was now going to leave his moorings ; he recommended the ladies to their couch and appointed an urchin.

(Passage from *Tregarthen Hall* by James Garland labeled as “domestic” with a probability of .79)

The classification further underscores the gendered and classed nature of domesticity. Passages featuring female characters engaged in household affairs or emotional reflection were more likely to receive high domesticity scores, reinforcing historical associations between women, domestic spaces, and structures of power. Consider again how, in example 9, Dorothea performs a kind of matrimonial, feminized relation in this dialogue with Casaubon, acting as both scholarly secretary and emotional support. Meanwhile, lower-class settings often exhibited a more ambiguous domesticity score, particularly in spaces where work and home life intersected. This suggests that domesticity is not

merely a spatial designation but also a socio-cultural construct shaped by class and gender expectations.

Finally, instances of unexpected domesticity, such as domestic-like interactions occurring on ships or characters finding moments of intimacy in liminal spaces, challenge rigid binaries between public and private spheres. The model's handling of these cases suggests that while domesticity is often anticipated in certain spaces, its presence can also surface in other places where characters engage in acts of care, reflection, or emotional connection.

7. Conclusion

Our approach to modeling domestic space in 19th-century English and Irish fiction provides new insights into both the concept of domesticity and computational approaches to analyzing literary settings. Our findings challenge conventional narratives that rigidly define domesticity by location, instead emphasizing the importance of activities and interactions that create domesticity in a variety of spaces within the novel. By moving beyond toponymic markers and incorporating non-traditional spaces, our model demonstrates the fluidity of domesticity and its dependence on relational and narrative cues.

The validation of our model against ground truth data reinforces its reliability while also highlighting areas of ambiguity, particularly in dialogue-heavy passages. This methodological approach addresses a critical gap in digital humanities research, offering a scalable way to analyze non-toponymic spaces computationally. In doing so, our study contributes to a new quantitative history of domestic space, revealing unexpected patterns in where and how domesticity is represented across 19th-century novels.

Ultimately, our results reveal the 19th-century novel not as a monolithic expression of gendered and classed domesticity, but as an evolving exploration of what domestic space could be. The strong language of domesticity captured by our model suggests that these novels were not merely reinforcing hegemonic ideals but experimenting with different forms of domestic representation. By rethinking domesticity through a computational lens, we uncover a more nuanced and dynamic portrayal of space, identity, and social structure in the literary imagination of the period.

In a more recent development, we have adapted our model to predict the probability of domestic space at the level of entire scene segments rather than fixed-length text windows. This shift toward longer, semantically coherent units allows for a more nuanced analysis of spatial setting in narrative. By identifying self-contained scenes that share consistent spatial features, the model opens new possibilities for detecting and classifying domestic environments at a finer granularity. Building on this foundation, we are now developing a more detailed spatial annotation framework that distinguishes between specific domestic spaces, such as kitchens, living rooms, and bedchambers, enabling systematic comparisons of how different types of rooms function across and within texts. This scene-based approach not only enhances spatial modeling but also lays the groundwork for richer literary analyses of space and setting in the future.

8. Data Availability

Data and Code can be found here: https://github.com/literarylab/jcls_domestic_space. It has been archived and is persistently available at: <https://doi.org/10.5281/zenodo.17219574>.

9. Acknowledgements

We thank Annie Lamar and our student assistants Sophie Schwarzhappel and Julia Gershon for the support in the annotation work. We also thank Kent Chang for his idea of splitting the recognition approach into two steps, which increased the performance of our model.

10. Author Contributions

Svenja Guhr: Project administration, Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft

Jessica Monaco: Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft

Alexander Sherman: Conceptualization, Methodology, Data curation, Formal analysis, Writing – original draft

Matt Warner: Conceptualization, Methodology, Data curation

Mark Algee-Hewitt: Project administration, Conceptualization, Methodology, Data curation, Writing – review & editing

References

- Armstrong, Nancy (1990). *Desire and Domestic Fiction: A Political History of the Novel*. Oxford University Press. [10.1093/oso/9780195061604.001.0001](https://doi.org/10.1093/oso/9780195061604.001.0001).
- Baledent, Anaëlle, Yann Mathet, Antoine Widlöcher, Christophe Couronne, and Jean-Luc Manguin (2022). “Validity, Agreement, Consensuality and Annotated Data Quality”. In: *Proceedings of the Thirteenth Language Resources and Evaluation Conference*. Ed. by Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Jan Odijk, and Stelios Piperidis, 2940–2948. <https://aclanthology.org/2022.lrec-1.315> (visited on 09/09/2025).
- Bamman, David (2021). *BookNLP*. <https://github.com/booknlp/booknlp> (visited on 09/09/2025).
- Bamman, David, Kent K. Chang, Lucy Li, and Naitian Zhou (2024). “On Classification with Large Language Models in Cultural Analytics”. In: *Proceedings of the Computational Humanities Research Conference 2024*. Ed. by Wouter Haverals, Marijn Koolen, and Laure Thompson, 494–527. <https://ceur-ws.org/Vol-3834/paper119.pdf> (visited on 09/09/2025).

- Bamman, David, Sejal Popat, and Sheng Shen (2019). "An Annotated Dataset of Literary Entities". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Ed. by Jill Burstein, Christy Doran, and Thamar Solorio. Association for Computational Linguistics, 2138–2144. [10.18653/v1/N19-1220](#).
- Bologna, Federica (2020). "A Computational Approach to Urban Space in Science Fiction". In: *Journal of Cultural Analytics* 5 (2). [10.22148/001c.18120](#).
- Cer, Daniel, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St. John, Noah Constant, Mario Guajardo-Cespedes, Steve Yuan, Chris Tar, Brian Strophe, and Ray Kurzweil (2018). "Universal Sentence Encoder for English". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Ed. by Eduardo Blanco and Wei Lu. Association for Computational Linguistics, 169–174. [10.18653/v1/D18-2029](#).
- Chadwyck-Healey Literature Collections and ProQuest, eds. (2016). *Nineteenth-Century British Fiction*.
- Cohen, Monica F. (2017). "Domesticity in Victorian Literature". In: *Oxford Research Encyclopedia of Literature*. Ed. by Deidre Shauna Lynch. Oxford University Press. [10.1093/acrefore/9780190201098.013.252](#).
- Davidoff, Leonore and Catherine Hall (1987). *Family Fortunes: Men and Women of the English Middle Class, 1780-1850*. Women in Culture and Society. University of Chicago Press.
- Dennerlein, Katrin, Thomas Schmidt, and Christian Wolff (2023). "Computational Emotion Classification for Genre Corpora of German Tragedies and Comedies from 17th to Early 19th Century". In: *Digital Scholarship in the Humanities* 38 (4), 1466–1481. [10.1093/llc/fqad046](#).
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Ed. by Jill Burstein, Christy Doran, and Thamar Solorio. Association for Computational Linguistics, 4171–4186. [10.18653/v1/N19-1423](#).
- Fludernik, Monika and Suzanne Keen (2014). "Introduction: Narrative Perspectives and Interior Spaces in Literature Before 1850". In: *Style* 48 (4), 453–460. [10.5325/style.48.4.453](#).
- Freedgood, Elaine (2006). *The Ideas in Things: Fugitive Meaning in the Victorian Novel*. University of Chicago Press.
- Frerichs, Sarah Cutts (1974). "Elizabeth Missing Sewell: A Minor Novelist's Search for the Via Media in the Education of Women in the Victorian Era". PhD thesis. Brown University.
- Google (2023a). *Experts/Bert. Version 2.0*. <https://www.kaggle.com/models/google/experts-bert> (visited on 09/09/2025).
- (2023b). *Universal-Sentence-Encoder. Version 2.0*. <https://www.kaggle.com/models/google/universal-sentence-encoder> (visited on 09/09/2025).
- Kababgi, Daniel, Giulia Grisot, Federico Pennino, and Berenike Herrmann (2024). "Recognising Non-named Spatial Entities in Literary Texts: A Novel Spatial Entities Classifier". In: *Proceedings of the Computational Humanities Research Conference*

2024. Ed. by Wouter Haverals, Marijn Koolen, and Laure Thompson, 472–481. <https://ceur-ws.org/Vol-3834/paper59.pdf> (visited on 09/09/2025).
- Kingma, Diederik P. and Jimmy Ba (2015). “Adam: A Method for Stochastic Optimization”. In: *Conference Track Proceedings of the 3rd International Conference on Learning Representations*. Ed. by Yoshua Bengio and Yann LeCun. [10.48550/arXiv.1412.6980](https://arxiv.org/abs/1412.6980).
- Krippendorff, Klaus (2018). *Content Analysis: An Introduction to Its Methodology*. 4th ed. SAGE.
- Lukács, Georg (1983). *The Historical Novel*. University of Nebraska Press.
- Marcus, Sharon, ed. (2007). *Between Women: Friendship, Desire, and Marriage in Victorian England*. Princeton University Press.
- Montani, Ines, Matthew Honnibal, Adriane Boyd, Sofie Van Landeghem, and Henning Peters (2023). *explosion/spaCy: v3.7.2: Fixes for APIs and requirements*. [10.5281/zenodo.10009823](https://zenodo.org/record/10009823).
- Moretti, Franco (1999). *Atlas of the European Novel, 1800-1900*. Verso.
- Parigini, Margherita and Mike Kestemont (2022). “The Roots of Doubt. Fine-tuning a BERT Model to Explore a Stylistic Phenomenon”. In: *Proceedings of the Computational Humanities Research Conference 2022*. Ed. by Folgert Karsdorp, Alie Lassche, and Kristoffer Nielbo, 72–91. https://ceur-ws.org/Vol-3290/long_paper399.pdf (visited on 09/09/2025).
- Piatti, Barbara (2016). “Mapping Fiction: The Theories, Tools and Potentials of Literary Cartography”. In: *Literary Mapping in the Digital Age*. Ed. by David Cooper, Christopher Donaldson, and Patricia Murrieta-Flores. Routledge, 88–101.
- Pichler, Axel and Nils Reiter (2022). “From Concepts to Texts and Back: Operationalization as a Core Activity of Digital Humanities”. In: *Journal of Cultural Analytics* 7 (4). [10.22148/001c.57195](https://doi.org/10.22148/001c.57195).
- Reiter, Nils (2020). “Anleitung zur Erstellung von Annotationsrichtlinien”. In: *Reflektierte algorithmische Textanalyse*. Ed. by Nils Reiter, Axel Pichler, and Jonas Kuhn. De Gruyter, 193–202. [10.1515/9783110693973-009](https://doi.org/10.1515/9783110693973-009).
- Ryan, Marie-Laure (2014). “Space”. In: *The Living Handbook of Narratology*. Ed. by Peter Hühn, John Pier, Wolf Schmid, and Jörg Schönert. <https://www-archiv.fdm.uni-hamburg.de/lhn/node/55.html> (visited on 09/09/2025).
- Ryan, Marie-Laure, Kenneth E. Foote, and Ma’oz ‘Azaryahu (2016). *Narrating Space/Spatializing Narrative: Where Narrative Theory and Geography Meet*. The Ohio State University Press.
- Schumacher, Mareike (2023). *Orte und Räume im Roman: Ein Beitrag zur digitalen Literaturwissenschaft*. J.B. Metzler. [10.1007/978-3-662-66035-5](https://doi.org/10.1007/978-3-662-66035-5).
- Schumacher, Mareike, Marie Flüh, and Marc Lemke (2022). “The Model of Choice. Using Pure CRF- and BERT-based Classifiers for Gender Annotation in German Fantasy Fiction”. In: *Digital Humanities 2022 Conference Abstracts*. DH2022 Local Organizing Committee. <https://dh2022.dhii.asia/dh2022bookofabsts.pdf> (visited on 09/09/2025).
- Soni, Sandeep, Amanpreet Sihra, Elizabeth Evans, Matthew Wilkens, and David Baman (2023). “Grounding Characters and Places in Narrative Text”. In: *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Ed. by Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki. Association for Computational Linguistics, 11723–11736. [10.18653/v1/2023.acl-long.655](https://arxiv.org/abs/1806.03268).

- SpaCy (2024). *en_core_web_sm-3.8.0. Models for English*. https://github.com/explosion/spacy-models/releases/tag/en_core_web_sm-3.8.0 (visited on 09/09/2025).
- Sutherland, John (1989). *The Stanford Companion to Victorian Fiction*. Stanford University Press.
- van Zundert, Joris J., Marijn Koolen, Julia Neugarten, Peter Boot, Willem van Hage, and Ole Mussmann (2022). “What Do We Talk about when We Talk about Topic?” In: *Proceedings of the Computational Humanities Research Conference 2022*. Ed. by Folgert Karsdorp, Alie Lassche, and Kristoffer Nielbo, 398–410. https://ceur-ws.org/Vol-3290/short_paper5533.pdf (visited on 09/09/2025).
- Wilkens, Matthew (2013). “The Geographic Imagination of Civil War-Era American Fiction”. In: *American Literary History* 25 (4), 803–840. <https://www.jstor.org/stable/43817603> (visited on 09/09/2025).
- Yang, Ziyi, Yinfei Yang, Daniel Cer, Jax Law, and Eric Darve (2021). “Universal Sentence Representations Learning with Conditional Masked Language Model”. In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Ed. by Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih. Association for Computational Linguistics, 6216–6228. [10.18653/v1/2021.emnlp-main.502](https://doi.org/10.18653/v1/2021.emnlp-main.502).