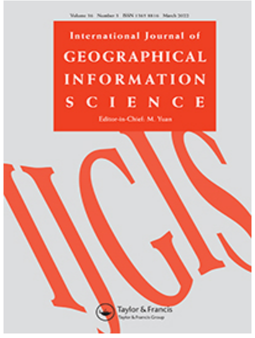

International Journal of Geographical Information Science

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tgis20

Detecting geospatial location descriptions in natural language text

Kristin Stock, Christopher B. Jones, Shaun Russell, Mansi Radke, Prarthana Das & Niloofar Aflaki

To cite this article: Kristin Stock, Christopher B. Jones, Shaun Russell, Mansi Radke, Prarthana Das & Niloofar Aflaki (2022) Detecting geospatial location descriptions in natural language text, International Journal of Geographical Information Science, 36:3, 547-584, DOI: 10.1080/13658816.2021.1987441

To link to this article: https://doi.org/10.1080/13658816.2021.1987441

Published online: 22 Dec 2021.

INTERNATIONAL JOURNAL OF GEOGRAPHICAL INFORMATION SCIENCE
2022, VOL. 36, NO. 3, 547–584
https://doi.org/10.1080/13658816.2021.1987441

RESEARCH ARTICLE

OPEN ACCESS

Check for updates

Detecting geospatial location descriptions in natural language text

Kristin Stock[a], Christopher B. Jones[b], Shaun Russell[a], Mansi Radke[c], Prarthana Das[c] and Niloofar Aflaki[a]

[a]Massey Geoinformatics Collaboratory, Massey University, Auckland, New Zealand; [b]School of Computer Science and Informatics, Cardiff University, Cardiff, UK; [c]Computer Science and Engineering Department, Visvesvaraya National Institute of Technology, Nagpur, India

ABSTRACT

References to geographic locations are common in text data sources including social media and web pages. They take different forms from simple place names to relative expressions that describe location through a spatial relationship to a reference object (e.g. *the house beside the Waikato River*). Often complex, multi-word phrases are employed (e.g. *the road and railway cross at right angles; the road in line with the canal*) where spatial relationships are communicated with various parts of speech including prepositions, verbs, adverbs and adjectives. We address the problem of automatically detecting relative geospatial location descriptions, which we define as those that include spatial relation terms referencing geographic objects, and distinguishing them from non-geographical descriptions of location (e.g. *the book on the table*). We experiment with several methods for automated classification of text expressions, using features for machine learning that include bag of words that detect distinctive words, word embeddings that encode meanings of words and manually identified language patterns that characterise geospatial expressions. Using three data sets created for this study, we find that ensemble and meta-classifier approaches, that variously combine predictions from several other classifiers with data features, provide the best F-measure of 0.90 for detecting geospatial expressions.

ARTICLE HISTORY
Received 18 August 2019
Accepted 26 September 2021

KEYWORDS
Geospatial language; natural language processing; spatial role labelling; geospatial parsing; spatial relations; locative expressions; georeferencing; geographic information retrieval

1. Introduction

References to geographic locations are common in text data sources including social media and web pages, and methods have been developed for their extraction and use through georeferencing of such resources. The georeferencing process is typically directed towards the detection of toponyms (place names) that serve as the basis of the resulting georeference. The default assumption is that the geographic reference is absolute, and thus one or several toponyms are treated as asserting the location or locations to which the document refers. However, many references to geographic locations are embedded in text without the presence of a toponym. Furthermore,

CONTACT Kristin Stock k.stock@massey.ac.nz

there is usually no consideration of qualifying phrases that describe the location of an object in space through its relationship to a reference object (e.g. *the church is beside the post office*).

It is important to take account of these qualifying phrases in georeferencing approaches because often a location might be described in a location description, or locative expression, as being *outside, some distance away from, to the north of*, or *in front of* (etc.) the named place, and thus refer to a quite different location than that georeferenced if only the place name is used. For example, the coordinates of Orewa would not describe the location referred to by the phrase *the accident occurred 30 km north of Orewa* very accurately. In general, such relative descriptions of location are rarely considered in georeferencing processes and then only with regard to a limited set of terms (Doherty *et al*. 2011, Hall *et al*. 2011, Chen *et al*. 2018).

To give another example, while the use of absolute location (i.e. a toponym) might often be reasonable when the objective is simply to determine an approximate geographic footprint for a document (Melo and Martins 2017), there are situations in which the use of the reference location in isolation will result in significant and potentially unacceptable errors in georeferencing. This may occur when precision in the location is required,for example, in the quoted phrase above, or when describing the location at which a sample of the natural environment was found, as in geology, biology, soil science or archaeology. Consider the description *samples were collected from the margin of Lake Vanda, in and around the north-eastern corner of the lake'* In georeferencing these samples it is important that the inferred georeference is actually at the north-eastern corner of the lake as the habitat there is significantly different from other parts of the lake margin, and certainly from the centre of the lake which is the location that would be provided by some gazetteer references to Lake Vanda.

The aim of our work is to detect the presence of expressions that contain relative geospatial location descriptions. This can be regarded as an essential pre-requisite to the development and application of methods to georeference their content. The proportion of geospatial location descriptions found in text is typically very low, with Stock *et al*. (2013) finding that only 0.2% of sentences in text across a range of geographically targeted web sources contained geospatial expressions, and thus pre-filtering of text to identify geospatial expressions using our presented methods may result in significant efficiencies in the georeferencing task.

Specifically, we address the challenge of developing automated methods for identification of geospatial locational expressions and distinguishing them from other-spatial (but not geographic) expressions and from non-spatial expressions.[1] Our definition of geospatial, other-spatial and non-spatial expressions (provided in detail in Section 2) depends in part upon the presence of a spatial relation term in the expression. By our definition, geospatial expressions include a spatial relation term and a reference object that is geographic, meaning that it is found outdoors or in transitional spaces and is unlikely to move in the normal course of events. Other-spatial expressions are those that include a spatial relation term but a reference object that is not geographic and may be small-scale, indoor or mobile. Non-spatial expressions are those that do not meet the criteria for either of the other two classes. See Table 1 for examples of geospatial, other-spatial and non-spatial expressions).

Table 1. Examples of geospatial, other-spatial and non-spatial texts from our collection.

Class of expression	Example
Geospatial expressions	*Between Sharps beach, and Angel's Beach is Flat Rock, a large rocky outcrop into the sea.* *Colwick Hill has its western part in the parish, and it is a fine relic of resistance to mighty floods of vast ages ago.* *A brick wall is reached beyond a bridge, with a number of obstructions over the next stretch.*
Other-spatial expressions	*A pin dropping in the attic would have shattered the silence.* *Alexandra looked uncomprehending and moved to put her spoon in her soup.* *After the battle the golden crown was taken from the helm of the fallen Plantagenet – the last of his line – and placed upon the head of the first of our Tudor kings.*
Non-spatial expressions	*He delighted his acolytes with tales about his time in the Army and how he had once lost an entire platoon.* *In the civil war of 324 he had represented his military campaign as a crusade against a corrupt paganism.* *In the film's most inventive touch, Hook tries to avenge himself on Peter Pan by becoming a second and better father.*

We define a spatial relation as the geometric configuration of two objects in space (relative to each other), with one object acting as a reference to describe the location of the other, located object (Carlson-Radvansky *et al*. 1999, Hudelot *et al*. 2008), and we refer to terms that describe that relation as spatial relation terms. Spatial relation terms could denote all typical configurations in space between the located and reference objects, such as those of proximity, coincidence, connectivity, containment, orientation and dynamic relations that express movement relative to a reference object. Spatial relation terms are often prepositions, but may also be verbs, adverbs, adjectives and other parts of speech (Dittrich *et al*. 2015). Table 2 provides some examples of spatial relation prepositions for different parts of speech, such as prepositions, including *near, at, between, of, surrounds, across, south of* and *outside*. Verbs can indicate spatial relations such as containment with *contains* and *has*, crossing with *cross*, the origin of a route with *departs* or the boundary of a linear feature with *lined*. Adverbs may be used in association with prepositions to qualify

Table 2. Examples of different parts of speech to describe spatial relations.

Spatial relation term part of speech	Example (spatial relation terms are shown in bold)
Preposition	*27kms* ***south of*** *Rotorua;* ***between*** *the shelled-out buildings;* *house built* ***at*** *Southwell; city block* ***in*** *Budapest; Library* ***on*** *Hucknall Road; Howie comes* ***from*** *Wormit; pilgrimage* ***to*** *Grantchester;* *the book is lying* ***on*** *the table;* ***outside*** *the Lithuanian capital;* ***round*** *the building; marshes* ***of*** *South Holderness;* ***near*** *Frome;* ***across*** *the border; the house was* ***in*** *the country;* ***behind*** *the waterfall*
Verb	*Clumber park* ***contains*** *about 4,000 acres; Abacus Airport Cars Cambridge* ***rides*** *there; street* ***lined*** *by tiny, dark shops;* ***departs*** *Mercure Grosvenor Hotel; Davies* ***left*** *Lansdowne Road;* ***Cross*** *the Europe-Asia divide*
Adverb/ adverbial phrase	*micro-distillery located . . .* ***just*** *off Oakland's waterfront; (main station) is* ***right*** *next door to the famous Tivoli Gardens;* *marched* ***southwards***
Adjective	***roadside*** *bomb;* ***southern*** *Italy;* ***south*** *side of the road;* ***shuttered*** *houses;* ***south-western*** *England;* ***tree-lined*** *road*
Possessive apostrophe (parthood)	*the* ***town's*** *Market Square;* ***South Side's*** *vertical farm;* ***Detroit's*** *Auto Museums;* ***Melbourne's*** *Dandenong Ranges;* ***Savannah's*** *Historic District;* ***Sydney's*** *Bondi Beach*
Mixed	***at the corner of*** *Main Street and West Broad Street;* ***Perth's riverside*** *resort;* ***on*** *a mountain road* ***in South of*** *France;* ***on the edge of*** *a housing estate;* *the canal* ***passing through*** *his land* ***at*** *Thornton;* ***On the eastern boundary of*** *the great forest;* *100 km* ***from*** *the busy port of Buenaventura,* ***west of*** *Bogota.*

spatial relations (***just*** *across the road;* ***directly*** *beside the house*) and as location indicators without an object (e.g. *marched* ***southwards***) (Garner 2017). Adjectives can be used, for example, to communicate spatial relations of bordering (a road) with *roadside*, the inclusion of a building of shutters with *shuttered* and the southern part of an object with *southern* and *south*, referred to as an internal relationship in Tenbrink (2011). Apostrophes can be used to indicate containment as, for example, in *Detroit's Auto Museums* and *Melbourne's Dandenong Ranges*.

The nature and characteristics of spatial language and its use to describe location have been discussed in a number of seminal works (Jackendoff 1983, Talmy 2000, Levinson 2003, Coventry and Garrod 2004), with, for example, Tyler and Evans (2003) describing the notion of spatial scenes, which 'involve conceptualizing a spatio-relational configuration between entities we encounter in the world around us and with which we interact' (p. 16). Herskovits (1987) focused on the notion of a locative as something indicating location or place and defined a *locative expression* as 'involving a locative prepositional phrase together with whatever the phrase modifies (noun, clause, etc.)' (Herskovits 1987, p. 7) and Kordjamshidi *et al*. (2017) distinguish between simple and complex locative statements, the latter with more than one landmark, and ask the question 'is there a spatial description in the sentence?' to identify locative statements. Herskovits introduced several forms of locative expression; the standard form consisting of a preposition and two noun phrases, as in *the house beside the river*, where the preposition *beside* describes the spatial relationship between the noun phrases, *the house* and *the river*. These two noun phrases have been described variously as the located object and the reference object, the trajector and landmark, the figure and ground and the locatum and relatum (Talmy 1983, Tenbrink 2011). In geospatial location expressions, the reference object is a geographical feature, while in other-spatial locational expressions, it is a feature at some other scale.

Our interpretation of geospatial language is broader than the locative expressions of Herskovits (1987) in that it does not prescribe which form of grammatical construction is used to communicate the spatial relation. As pointed out above and in Table 2, spatial relations may be described using several parts of speech. Notably, clear, unequivocal and specific definitions of what counts as spatial language are scarce, and when we consider the many possible ways in which spatial relations can be expressed, there are many difficult, borderline cases. The challenges become even more complex when we attempt to define geospatial language and distinguish it from other-spatial languages (Gritta *et al*. 2018, Wallgrün *et al*. 2018).

It should be noted that we do not have explicit rules to include or exclude particular grammatical forms or specific spatial relation terms. This stems from the fact that, as illustrated in Table 1, there are many terms, such as *in*, that can be used in multiple contexts with different meanings that may or may not communicate spatial information. We use statistical methods to learn from examples of locational expressions, where these methods take account of all words in the expression (not just the spatial relation terms) to make a decision as to whether it is geospatial, other-spatial or non-spatial.

Our definition of spatial relations provided above is different from the spatial relations referred to in the context of geographical information systems and spatial databases in which spatial relation query operators such as *inside, overlap* and *touch* have a formal geometric definition that allows them to be implemented unambiguously. In contrast, we are concerned with detecting expressions that contain linguistic spatial relation terms that may have multiple senses, varying in meaning by context and encompassing vagueness and ambiguity. Another alternative meaning for the phrase *spatial relation* can be found in the field of natural language processing, in which it refers to a specific form of semantic relationship between a pair of entities (Surdeanu *et al*. 2012) where that relationship is spatial (Zhang *et al*. 2009). The phrase refers to not just the spatial relation terms but also all components of the relationship and our use of the *spatial relation term* is similar to that of *the spatial indicator* in (Kordjamshidi *et al*. 2011).

Our research questions can be summarised as follows:

(1) Can we develop automated methods to identify natural language expressions that describe the relative location of geographical phenomena and distinguish these from other-spatial descriptions of phenomena such as smaller indoor or table-top scales and from non-spatial expressions?
(2) What patterns of speech and terminology distinguish geospatial locational expressions, other-spatial locational expressions and non-spatial expressions?

It should be noted that in answering these questions, our objective is to determine whether a sentence includes expressions that describe locations that are either geospatial or other-
spatial not to extract the individual components such as the located and reference objects.

In addressing the research questions, we experimented with four approaches to the creation of features for machine learning classifiers designed to recognise the presence of geospatial expressions and distinguish them from other-spatial and non-spatial expressions. In the first approach, we use a bag of words method in which each word in the vocabulary of all sentences is treated as a feature for machine learning. In the second approach, we use the word embeddings of each word of an expression to generate an averaged embedding to represent the entire sentence. The bag of words approach can be regarded as representing the meaning of a sentence in terms of the presence or absence of words present in the document collection, while word embeddings represent the meaning of individual words with multi-dimensional vectors, created with a dimensionality reduction procedure that operates on the association between the embedded word and other words with which it co-occurs. In the third approach, the features are the frequency of occurrence of various manually generated syntactic language patterns that characterize geo-locational expressions, combining individual elements (for example, the combination of a spatial preposition and a geographic object type). As a baseline, we use only the presence of a toponym as a feature in a classifier. In addition, we experiment with several ensemble or meta-classifiers that combine predictions from the bag of words, word embedding and pattern classifiers with, variously, the word embeddings, the language pattern features and the place name features. Evaluation of the methods was conducted with three test datasets created for this study. This includes an analysis of the predictive power of particular words and language patterns in distinguishing the different types of expressions in general and with reference to example sentences.

In the following section, we define our interpretation of geospatial locational expressions, other-spatial locational expressions and non-spatial expressions and explain with further examples of some of the multiple ways in which geographical locations can be described relative to reference objects. This is followed in Section 3 with a review of related work regarding automated methods that assist in the process of identifying and extracting content from such sentences. In Section 4 we describe our test data sets that have come from several sources. Our various machine learning methods for detecting locational expressions are explained in detail in Section 5, while in Section 6, we present the results of the experimental evaluation. Conclusions and a discussion of future work are provided in Section 7.

2. Classifying expressions as geospatial, other-spatial or non-spatial

Here, we classify natural language expressions into three mutually exclusive classes of geospatial language, other-spatial language (language that is spatial but not geospatial) and non-spatial language. Although our main focus is geospatial language, consideration of other-spatial language enables comparison of our methods with generic spatial classification schemes (that combine both geospatial and other-spatial expressions) and a finer grained analysis of language content.

2.1. Geospatial locational expressions

To identify relative geospatial expressions that are candidates for being georeferenced, we define geospatial expressions as having the following characteristics:

(i) They include a spatial relation term, which may take the form of a preposition, or a verb or other words or a group of words that describe the spatial location or movement of one object relative to another (see Table 2 for examples using different parts of speech).
(ii) The reference object (also known as a landmark or relatum) of the spatial relation is a geographical object. By geographical object, we mean an object that is found outdoors or in transitional spaces that are large and public (Kray *et al*. 2013); that is static in nature and unlikely to move in the normal course of events.

Geographical objects are typically the kinds of objects that are likely to occur on a map, ranging in scale from street furniture (lamp posts, fire hydrants), such as might appear on an engineering or landscape map and on some detailed topographic maps, up to objects on a global scale and thus encompasses some of the objects found in Montello's vista space, as well as those in his environmental and geographic spaces (Montello 1993). Note that the reference object is usually represented grammatically either by a proper noun or a noun and could be either a named place (toponym) or a type of geographic object.

We define geospatial language in this way and distinguish it from other kinds of spatial language, in order to identify those expressions that we can map or for which we can determine the geographic location, as this is important for a number of applications (e.g. location-based services and geographic information systems). Such georeferencing may be direct (e.g. if a place name is mentioned) or indirect (e.g. if a geographic object type is mentioned). We include the latter case because it would be possible to georeference such expressions in one of the two ways. First, the reference object may be tied to a geographic location through a coreference. For example, the text below includes a specific place name for the reference object (the river), but it is not in the same sentence as the phrase *hotel by the river*. We cannot georeference *hotel by the river* in isolation, but by resolving the coreference to the previous sentence, georeferencing is possible.

> *The Trent River runs through the centre of Nottingham and is frequently used by canal boats. When we visited in July, we stayed in a hotel by the river.*

Second, location expressions that refer to geographic feature types may be used to georeference groups of features that meet specific criteria using a spatial query. For example, candidate locations for the expression *hotel beside train station* could be found by identifying multiple train stations that have hotels beside them within a known broader geographic area (e.g. a specific city), as recorded in geographic data sets containing train station and hotel locations.

Our definition of geospatial does not include expressions that contain only a place name, with no spatial relation term. The reason for this is that research into Named Entity Recognition, to detect place names in isolation, is already well-developed (Won *et al.* 2018). Our contribution is the detection of more complex references to location.

Examples of geospatial expressions include the following:

(1) *Maybe, thousands of years ago, birds and reptiles from continental South America had reached the Galapagos, ferried on the rafts of vegetation that float down the rivers and out to sea.*
(2) *About 100 homes across Te Puke, Omanu, Matapihi, and Paengaroa were without power on the night of the 19th.*
(3) *Climb the hill and enter the echoing fifteenth-century Gothic church to peer through glass panels at the medieval foundations.*
(4) *This fort is in the centre of the city near the sea-front and in front of it is a very pleasant tree-lined road with an open-air café.*
(5) *All over the United States, people are fleeing urban areas with high infection rates for the perceived safety and natural beauty of rural areas*

The first two examples include spatial relation terms and named geographical locations and thus could be directly resolved to a specific geographic location. In Example 1, there are a number of spatial relation terms including prepositions (*from, out to, on and down*) and the verb *reached* to describe the path of birds and reptiles. The spatial relation term associated with South America is *from* meaning was *located at*, while the spatial relation term associated with the Galapagos of *reached* also conveys having been *located at*. In Example 2 the spatial relation term *across* is a preposition and applies to a list of named geographic entities. There is also a geographic feature type *homes*. The third example has no named places that would occur in a gazetteer, but it contains several geographic feature types (*hill, church* and *foundations*) with the spatial preposition *at* and with verbs of motion that convey dynamic spatial relations of moving across a surface (*climb*) and crossing the boundary (*enter*) of the respective reference objects. The preposition *through* is also a spatial relation term but with a non-geospatial object (*glass panels*). The fourth example contains several geographic feature types (*fort, city, sea-front, road* and *café*) associated with the spatial relation prepositions *in, the centre of, near, in front of* and *with (*meaning proximity or containment). We include this kind of expression in our definition as an example of the use of multiple reference locations that in this case are not named. Resolution of such expressions would require additional geographic data about locations of forts, cities, etc., in the country concerned (e.g. from the wider discourse) and could then be resolved using spatial queries. Also, in this example, the adjective *lined* is used to express the spatial relation of adjacency between the road and the trees. The final example includes three spatial relation terms: *all over* (*people all over the United States*); *fleeing*, meaning moving away from (*people fleeing urban areas*), and *for*, meaning towards (as in *for rural areas*). The first of these can be georeferenced using the known location of the United States. The second and third could be combined with existing geographic data sets of urban and rural areas to provide an indication of possible migration patterns.

2.2. Other-spatial locational expressions

Other-spatial expressions are defined here by extension from geospatial expressions. Like geospatial expressions, they contain a spatial relation term using the same definition as for geospatial expressions (item *i* above). However, for other-spatial expressions, the reference object of the spatial relation term does not meet the criteria for geospatial expressions (item *ii* above), in that it could be indoor, mobile or small in scale, as in table-top space or in the figural space of Montello (1993).

Examples of other-spatial expressions include:

(1) *Masklin darted between two seats, around a pair of giant shoes, and threw himself flat on the carpet.*
(2) *Reaching for a napkin, Ashley tucked it beneath Thomas's chin.*
(3) *Frank stood up and, taking up an indignant posture, he placed one hand on his hip and pranced towards her, saying, ` No alcohol ever crosses my lips, apart from cooking sherry'.*

All three of these expressions include spatial relation terms, in the form of prepositions *(between, around, beneath*) and verb-preposition combinations (*reaching for*, meaning moving toward), but the reference objects do not qualify as geographic objects due to scale and/or mobility (*seats, napkin, chin, nose and hip*).

2.3. Non-spatial expressions

Non-spatial expressions are those that do not fall into either of the above classes in that they do not have both a spatial relation term and a reference object that refers to one or more specific real-world objects.

Examples of non-spatial expressions include the following:

(1) *He's even been stealing from the business to get cash to buy his drugs.*
(2) *You look as if you've been bickering with a volcano.*

Neither of these expressions includes a spatial relation term. The first refers to a criminal financial operation that does not clearly involve any specific spatial movement of cash. In the second expression, the use of volcano as a reference object is metaphoric and does not refer to an object that could be georeferenced.

The task of distinguishing geospatial, other-spatial and non-spatial expressions is not straightforward, as often the same term may be used in each expression to communicate a geo-spatial relation, an other-spatial relation and a non-spatial relation. For example, *she is outside the post office; the book is outside the box; it's outside my expertise* all use the term *outside*, but taking account of their reference objects they could be classed, respectively, as geospatial (referring here to a static reference object in geographic space), other-spatial (referring here to a movable reference object in table-top space) and non-spatial (because the use of *outside* is metaphoric and *expertise* is not an object that can be spatially referenced). It may be noted that while other studies have developed methods for detecting general locational expressions, which may or may not include an explicit place name, to our knowledge, none has distinguished geospatial from other-spatial locational expressions. Some previous work has, however, focussed on distinguishing generic spatial language from non-spatial language at the level of individual prepositions (Kordjamshidi *et al*. 2011) and the problem of identifying prepositions that can be classed as geospatial was addressed in Radke *et al*. (2019).

While the similarities between geospatial and other-spatial location descriptions are significant, the distinction between the two is important. Of particular importance, geospatial language can be georeferenced using a geographic reference system (for example, latitude and longitude) that ties a location to the Earth's surface. This is in contrast to the other most common approaches to interpretation and generation of spatial language, which typically reference a local indoor or desktop reference frame, notably in robotics applications (Kunze *et al*. 2014, Spranger *et al*. 2016, Paul *et al*. 2018). In principle, it may be possible to link indoor or table-top space to geographic space, but descriptions of indoor or table-top environments cannot normally be directly georeferenced either through the geocoding of the reference objects or of terms that co-reference such objects. Examples of non-geographic reference objects include eating utensils (*the knife is next to the fork*); furniture (*the book lies on the table*); and body parts (*the necklace was around her neck*).

3. Related work on automated detection of spatial language

The task of detecting geospatial language can be regarded as analogous to work on spatial relation labelling and extraction, for which several methods have been developed, but that work (summarized below) is largely focused on a subset of types of spatial relation terms (e.g. prepositions) and does not distinguish explicitly between geospatial and other forms of spatial relationships.

A rule-based system for spatial relation (spatial relation term, locatum and relatum) extraction for the Chinese language was presented in Zhang *et al*. (2009). They used a set of manually defined language patterns to extract the relations and report F1 values between 0.59 and 0.75 for different types of spatial relation terms. Kordjamshidi *et al*. (2011) achieve better performance for some aspects of spatial relation extraction (which they refer to as spatial role labelling) in English, although their work is confined to cases in which the spatial relation term, referred to as a spatial indicator, is a preposition. While as noted above, this is often the case, a spatial relationship can also be conveyed through other parts of speech (Table 2). Kordjamshidi *et al*. (2011) used machine learning methods in which, in the case of the spatial indicator (spatial relation term), the features were words dependent on the spatial indicator or on which the spatial indicator was dependent, along with parser-derived information such as the part of speech, the lemma of the term, the dependency relation and the semantic role (as determined by a semantic role labelling parser). For trajector (located object) and landmark (reference object) identification, additional information on the path in the parse tree, between the candidate word and a spatial indicator, was included. This paper presents two approaches to extracting the three components. The first is a pipeline in which a Naïve Bayes classifier is used to detect whether a preposition is being used with a spatial sense, before using a conditional random field (CRF) classifier to identify the trajector and landmark associated with the proposition. The second approach uses a joint learning method in which all three are found simultaneously using a CRF classifier. Our work may be regarded as analogous to the latter study with regard to the determination of the sense of spatial relation terms (in their case prepositions), but it differs significantly in that we are concerned with determining geospatial rather than generic spatial sense; we are concerned with the sense of an entire expression or sentence rather an individual preposition; and we do not confine our attention to prepositions but consider all forms of geospatial expression that use relative terminology (as opposed to simply listing a place name).

Hassani and Lee (2017) improved on Kordjamshidi's results in the detection of the generic spatial sense of prepositions by using a deep learning approach that combined word embeddings of the words in the local context window surrounding the preposition to be classified with lexical, syntactic and semantic features (such as a word and its lemma, parts of speech and named entity types and dependencies). As one of the methods in our work, we also use word embeddings as features, in this case of all the words in the expression to be classified, but as indicated above, we are concerned with the geospatial sense of whole expressions rather than the generic sense of individual prepositions.

A study of the adaptation of spatial role labelling methods to the task of detecting whether a preposition has a specific geospatial sense (as opposed to the generic spatial sense) was conducted by Radke *et al*. (2019), but the performance of their system was limited to a best F1 value of 0.64. Our task differs from the latter approach in that we are concerned with classifying entire expressions and we do not focus specifically on the sense of individual prepositions.

A machine learning method for identifying partial locative expressions that consist of a preposition (that specifies the type of spatial relation) and a reference object is presented in Liu (2013). They refer to these expressions as degenerate locative expressions (DLEs). Similar to Kordjamshidi *et al*. (2011), Liu uses natural language processing features such as the word itself, part of speech tag and lemma, as well as text chunking labels (e.g. the start of a noun phrase), the offset of the word in the sentence, the presence of any of a set of manually determined location indicative words (including nouns and verbs), and the geographic object type where it can be ascertained for toponyms. They do not use the dependency relations and semantic role information employed in Kordjamshidi *et al*. (2011). The method identifies the beginning and interior components of their locative expressions and, when applied in a fully automated mode to the

TellUsWhere corpus, obtained an F1 score of 0.77. The TellUsWhere corpus was obtained from a mobile game, in which participants were asked to describe their location, and consists largely of locative expressions. Their classifier does not attempt to distinguish between expressions with a geographical reference location and others such as indoor descriptions. This system was evaluated in Liu *et al*. (2014) in comparison to alternative approaches to detecting locative expressions and with some geo-parsers that only identify geographic named entities. When applied to a corpus of several sources of social media (including Twitter, Wikipedia and some blogs) along with the British National Corpus (BNC), it obtained an F1 score not higher than 0.16 (with the StanfordNER parser providing the best results going up to an F1 of 0.41). The low score was attributed to its having been trained on the TellUsWhere corpus (on which it achieved the F1 score of 0.77). Our work is only partially comparable to these methods in that we are not concerned with identifying the specific components of locational expressions, but focus on detecting whether a sentence includes a geospatial relational expression, or another form of spatial expression, or neither.

Khan *et al*. (2013) describe a method for extracting geospatial triples of <located object> <spatial indicator> <reference object> that uses the method of Liu (2013) to retrieve the spatial indicator and reference object. The rule-based approach uses the Stanford Parser to find prepositional phrases of the form <governor > <preposition> <dependent>. Where the preposition was equivalent to the spatial indicator of the extracted degenerate locative expression (DLE), the <governor> from the parser then provides the located object of a geospatial triple. The methods include other rules to add qualifying words to nouns, adjectives and verbs that are part of the DLE, hence enhancing the form, as well as a manual option to detect the place names that serve as the reference object. The method was applied to the TellUsWhere corpus. The authors distinguish partial DLEs from locative DLEs where the latter use explicit spatial relational terms such as *near* and *in*, while the former contain only prepositions such as *to* and *from* that can be enhanced to a spatial form when used for example as part of *next to* or *3rd house from*.

A rule-based approach to detecting toponyms in a reference sense, i.e. associated with a spatial relation term, was presented in Wolf *et al*. (2014) as part of a study to distinguish different uses of toponyms (including rhetorical forms). The rules used part of speech tags and a dependency parser to identify prepositional phrases and gave performance on a German language corpus that equates to an F1 value of 0.91. It may be noted that, as mentioned above, in our work, an expression could be classed as geospatial without including a toponym, provided that there is some form of geo-spatial reference object.

The detection of 'localization relationships' is a key aspect of Kordjamshidi *et al*. (2015) who present a specialised application of spatial role labelling methods to detect associations between bacteria, that are treated as trajectors, and habitats that are treated as landmarks. The approach is notable for using a machine learning method that employs a structured SVM (support vector machine) classifier in combination with multiple predominantly linguistic features that include the presence of both prepositions and verbs as indicators of possible spatial links between the associated entities.

Several studies have addressed the problem of identifying the presence of locational expressions in the context of building a corpus of geospatial expressions. Stock *et al*. (2013) describe methods to determine whether a sentence contains a geospatial relational expression, which is close in motivation to the present paper. They use a set of search patterns to retrieve candidate sentences from the Web before searching these sentences for the presence of various language patterns that are representative of geospatial location descriptions, achieving a precision of 0.66. Here, we evaluate the use of similar language patterns as features for machine learning in one of our approaches. In another project to create a corpus of spatial relational expressions, Wallgrün *et al*. (2014b) used web search engine query patterns to retrieve documents that were constrained to specific object types for the located and reference objects and to the three spatial relation terms of *near, close* and *next to*. The results were, however, validated manually.

4. Creation of the test data sets

In order to test the classification methods, we created three test data sets. The main data set, which we refer to as MT6.5K (as it was created with Mechanical Turk and contains 6579 expressions), consists of expressions from a number of different sources from a range of domains, sourced via The Preposition Project[2] (TPP) and the Nottingham Corpus of Geospatial Language (NCGL) (Stock *et al*. 2013). This data set was engineered in a way that ensured a balance of expressions in each of the three target classes, as described below. We then further test the methods on two domain-specific data sets, each containing 1000 expressions. The first (NIWA) contains reports of impacts from storm events, and the second (COVID) is web-scraped content about the COVID-19 pandemic.

We aim to classify natural language expressions into the three classes that were defined in Section 2. When applying the scheme to expressions that contain both geospatial and spatial elements, the geospatial classification takes precedence over the other-spatial classification. The definitions of the three classes in Section 2 form the basis of the explanation of the classes given to Mechanical Turk workers, documented in Appendix A.

4.1. MT6.5K data set

In order to train and evaluate the automated classifiers, we created a classified data set from a combination of two sources and using two different methods. The sources were as follows:

(1) The TPP data set used for The Preposition Project and SemEval-2007 (Litkowski and Hargraves 2007). The entire data set contains 24,413 sentences. While the ground truth senses of the prepositions provided with the data set might be used to infer generic spatial and non-spatial sense, we classified the expressions again using Mechanical Turk (described below) to comply with the class definitions in Section 2, in order to make the distinction between geospatial, other-spatial and non-spatial senses of the sentences.

(2) Data collected during the process of creation of the Nottingham Corpus of Geospatial Language[3] (NCGL), as described in Stock *et al*. (2013). The process of creation of the NCGL involved an automated first step followed by manual verification. Two groups of expressions were used in our test collection:
   (a) those that had been automatically harvested and manually verified and confirmed to be geospatial and were hence part of the final NCGL and
   (b) those that had been automatically harvested from the original sources from which the NCGL was derived, but whose class had not been manually checked and as such were not included in the NCGL.

We used Mechanical Turk[4] to classify the expressions from sources 1 and 2b as described in Section 4.2. Expressions from 2a had already been classified manually and confirmed to be geospatial, and a subset of these were included to balance the mix of different classes resulting from the Mechanical Turk process.

The challenges of objectively classifying location expressions have been well-recognized. For example, Wallgrün *et al*. (2018) present an approach to the creation of an annotated corpus identifying the presence and annotation of place names, consisting of two steps: first identifying place names, for which they use Mechanical Turk, and second disambiguating those place names (identifying the real-world location they refer to), for which they use geographic experts. They conclude that 'no place reference corpus can be perfect due to the nature of language; different individuals are likely to interpret the same text differently, and in some cases, the individual who generated the text may not even agree with themselves on its meaning at some later point in time' (p. 25). The challenges involved in identifying place names using a crowd-sourcing approach are also discussed in Clematide *et al*. (2018), who found that only 7 of 46 place names were identified by more than 50% of the crowd participants. Moving beyond place name identification to other elements of geospatial language, Aflaki *et al*. (2018) discuss the difficulty in achieving agreement between trained and paid non-experts who were given the task of identifying relata, locata and spatial relation terms, among other less common spatial word categories. The best inter-annotator agreement achieved was 0.65, for spatial relation terms.

This previous work demonstrates the challenge of involving non-experts in annotation tasks related to spatial and geospatial language. Our annotation task differs from this previous work in that we are addressing the task of annotation of language class rather than specific elements within expressions. Nevertheless, the task of annotating whether an expression is geospatial, other-spatial or non-spatial presents a considerable amount of ambiguity (see Section 2). This can give rise to what Wong and Lee (2013) term 'legitimate disagreement' among annotators. They refer to situations in which there is inherent ambiguity, as, for example, in word senses related to emotions or social acts, and propose annotating with a confidence score that, for multiple annotators, could reflect the proportion supporting an interpretation. They point out that for machine learning where more definite decisions are required, the annotations can be used to filter out more ambiguous instances. Here, we adopt a similar approach in that we omit individual cases where annotation agreement is weak to the extent that there is no clear majority in favour of any particular interpretation. It should be remarked, however, that such cases remain of interest in representing language that has a level of ambiguity and can be the focus of further study (Wallgrün *et al*. 2014a).

4.2. Method for classification of MT6.5K data set

We recruited a collection of Mechanical Turk workers to classify 8500 expressions, 2/3 of which came from source 1 above (TPP) and 1/3 from source 2b above (NCGL surplus) by posting a set of human intelligence tasks (HITs) on the Mechanical Turk site for workers to select. This split was used as we expected that source 2b would have a greater proportion of geospatial expressions than the other classes, given that its contents were harvested using a process that automatically selected geospatial expressions with 66% reliability (Stock *et al*. 2013). We collected 8 classifications for each expression (i.e. 8 workers classified each expression) and a total of 157 workers classified expressions. The number of expressions classified by each worker varied (maximum 6099, minimum 1), as workers had the option to complete as many or as few HITs as they wished. Workers were paid US $0.05 for each HIT, and one HIT involved classifying one individual expression. Expressions were no longer than one sentence. The instructions given to workers are contained in Appendix A. The descriptions of the classes were intentionally kept brief with only a few key examples, to ensure that respondents read them closely and paid attention to the details, but additional examples were provided via a link to web page (Appendix B).

The class for a given expression was determined as the mode (most frequently occurring) class given by the Mechanical Turk workers. However, the data was processed to remove poorly performing workers, and expressions with poor agreement, as follows.

Poorly performing workers were excluded, as defined by their degree of conformity with the mode classification across all expressions they classified. The conformity score for each worker *w* was calculated as follows:

$$conformity\ score_w = \frac{n_{conforming}}{number\ of\ expressions\ classified\ by\ w}$$

where $n_{conforming}$ = the number of expressions for which the class given by worker *w* agrees with the mode class for the expression where the mode was calculated across all 8 responses. Workers with conformity scores less than 0.7 were excluded, and the final classification of the expression was determined to be the modal classification of the remaining workers. Of the original 157 workers, 110 achieved the required conformity standard, and the classifications of the remaining 47 were excluded from further analysis. The conformity threshold of 0.7 was selected because it resulted in the highest level of average agreement across all expressions, and because higher thresholds made it difficult to achieve sufficient classifications for some expressions. For example, a conformity threshold of 0.7 results in an average of 5.8 workers classifying each expression, while a conformity threshold of 0.8 results in an average of only 4.3 workers classifying each expression.

Expressions for which there was poor agreement among workers were also excluded, following a similar approach to Potthast (2010) and Wallgrün *et al*. (2018). The agreement score for each expression *x* was calculated as follows:

$$agreement\ score_x = \frac{n_{agreeing}}{number\ classifications\ for\ x}$$

where *x* = expression and $n_{agreeing}$ = the number of classifications of x that are the modal class for the expression. Expressions with an agreement score of less than 0.8 were excluded, resulting in 5,664 of the original 8,500 meeting the required agreement threshold. The 0.8 threshold was chosen in order to achieve a balance between the number of expressions that were excluded and data quality. A threshold of 0.9 excluded all but 3688 expressions, which would have made subsequent training and evaluation of the methods more difficult. We acknowledge that this process of filtering out expressions that have poor agreement may result in the most contentious cases being excluded, but the nature of geospatial language suggests that universal agreement among humans is difficult to achieve (Wallgrün *et al*. 2018) and thus, it would be difficult for us to accurately validate machine classifications on these more contentious cases. As a final step, one of the paper authors conducted a blind (without access to the classes determined by Mechanical Turk) manual classification of a sample of 100 expressions randomly selected from the set of 5,664. The inter-annotator agreement (Cohen's kappa) between the two classifications was 0.95, indicating a very high level of agreement.

Table 3 shows the number of expressions of each class in the resulting data set. In order to achieve a more balanced data set for method evaluation, an additional 915 previously manually annotated geospatial expressions were added from source 2a above (see Section 4.1), giving equal numbers of geospatial and non-spatial expressions. This is because naturally occurring data sets (e.g. the NIWA and COVID data sets explained in Sections 4.3 and 4.4) often have very skewed numbers of expressions, which distorts the results and can make effective training difficult. Although the MT6.5K data set is drawn from multiple real data sources, it was curated in order to make it suitable for method development and comparison. We therefore included an additional two data sets in our evaluations, to evaluate the success of the methods with data 'in the wild'.

4.3. NIWA data set

The NIWA data set contains expressions that were extracted from the New Zealand Historic Weather Events Catalogue, hosted by the New Zealand National Institute for Water and Atmospheric Research (known as NIWA). The catalogue contains reports on major weather events and their impacts (damage, casualties, etc.) over the last 200 years, extracted from a range of publications including newspapers, journals and databases. We used the web portal[5] to search for all events that involved the hazard 'lightning' by selecting from the drop-down list provided. Since disaster reports are dominated by geospatial expressions, we selected the geographically limited hazard type of lightning in order to increase the likelihood of other-spatial expressions (due to impacts in small scale, indoor and personal space) in an attempt to maximise balance among classes. The resulting records were then downloaded in XML, and the contents of the 'impacts' tag extracted, again in order to identify precise impacts of the lightning events that might refer to specific geographic and other-spatial locations. We sentence tokenised (segmented into sentences) the impact

Table 3. Number of expressions in data sets.

	MT6.5K			
Class	Number from Mechanical Turk	Final Number Used	NIWA	COVID
Geospatial	1661	2576	695	465
Other-spatial	1427	1427	113	55
Non-spatial	2576	2576	192	480
Total	**5664**	**6579**	**1000**	**1000**

descriptions, which resulted in a data set of 15,895 sentences. We then selected 1000 sentences of these and manually labelled them. The selection of the 1000 expressions was triaged using a bag of words SVM classifier, training on the Mechanical Turk 6.5 K dataset and applying the model to the new NIWA lightning data set (15,895). We then randomly selected a sample from each of the three classes to produce a sample of 1000. The selection of 1000 was manually labelled by one author, and a random sample of 400 of those were labelled by another author, with an inter-annotator agreement accuracy of 0.91 and Cohen's kappa of 0.78. Any disagreements, together with expressions from the remaining 600 expressions that were flagged as ambiguous by the first author, were examined by the two labelling authors together to reach agreement.

The NIWA data set contained a number of expressions whose reference objects were weather events (e.g. *The winds that accompanied the waterspouts were described as horrendous; Several small water spouts were whipped up by the winds but fizzled out before developing into damaging tornadoes*.). While these meet some criteria of our geospatial class (they are in geographic space), due to their mobility and the absence of static reference objects, we classified them as other-spatial. As shown in Table 3, the data set is highly skewed towards geospatial expressions. The dominance of geospatial expressions in this data set is not surprising, given the subject matter of disaster event impacts, but given that many other data sets have a very different balance of the three classes, often with very low proportions of geospatial expressions, we also tested with a third data set.

4.4. COVID data set

The COVID data set was created by scraping web pages using the BootCat[6] tool, with seed words: *COVID-19, coronavirus, COVID, spread, city, town, country* and *distribution*. BootCat generates a set of triples from this set of seeds, then conducts web searches using those triples, and harvests the contents of the pages that are returned by the search. We sentence tokenized the results from the BootCat process, resulting in a data set of 3731 expressions. We used a similar triage and manual labelling process as for the NIWA Lightning data set, achieving an inter-annotator agreement accuracy of 0.89 and Cohen's kappa of 0.78. The final, manually labelled data set contained an approximately equal number of geospatial and non-spatial expressions, but only a small number of other-spatial (Table 3).

The COVID data set was extracted in September 2020, when large case numbers and community restrictions were active in some areas. As a result, the data set was characterised by a high occurrence of hypothetical and instructional expressions (e.g. *If you do not wish to return your ballot by mail, you may use the drop boxes that are located at the Town Hall; If you are not feeling well, stay home and get tested)* or very general expressions *(e.g. BOSTON [CBS] Massachusetts on Wednesday released its weekly report on the latest coronavirus case numbers by cities and towns*), and frequent references to official organisations (e.g. *The Town of Norwood in partnership with the South Middlesex Opportunity Council (SMOC) is pleased to announce a new grant program*) with implied locations, yet referring to an organisation, rather than a location. This sometimes presented challenges when assessing whether an expression was geospatial or not, as discussed in Section 6.

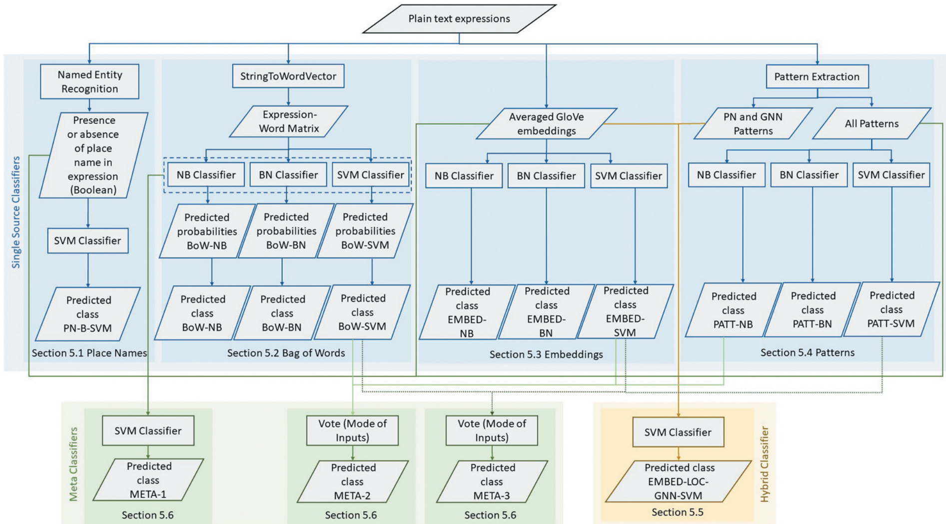

Figure 1. Classification methods overview.

5. Methods for detecting location descriptions

Several classification methods were developed for the purpose of determining automatically whether a natural language expression contains a relative description of location. We create ternary classifiers where the target class can take the value of geospatial, other-spatial or non-spatial. Each classifier was implemented using the Naïve Bayes, Bayesian Network and SVM methods. As indicated above, we consider four approaches to the generation of features, using a bag of words, word embeddings and language patterns, as well as a simple baseline that uses the presence or absence of place names. The meta-classifier approaches combine the features and the probability predictions provided by the other methods in various ways as shown in Figure 1. Each of the methods is described in more detail in this Section, and then their specific implementation for evaluation purposes is described in Section 6.

5.1. Place name only [PN]

This baseline method takes a naïve approach, classifying an expression based solely on the presence or absence of a place name, relying only on the Stanford NER tool to detect place names. We used the Stanford NER tool because in comparative studies such as Karimzadeh *et al.* (2019), the Stanford parser has been shown to be a very effective method of identifying locations and in Wang and Hu (2019) the system that used Stanford for NER outperformed the system that used SpaCy NER. The assumption is that if an expression contains a place name, this is an indication that it is geospatial. This method cannot therefore be expected to distinguish other-spatial and non-spatial from each other.

5.2. Bag of words [BoW]

This approach applies a bag of words classification to create a normalised matrix of expressions *vs* words, with each cell populated by the *tf-idf* value for the word in the respective expression (Luhn 1957, Salton and Buckley 1988). The width of the matrix is given by the vocabulary of words from the test collection and the features for each expression are therefore the *tf-idf* values for each of the entire vocabulary of these words (most values are zero as positive values will only be recorded for the words that are present in the respective expression). We used the most frequently occurring 1000 words, although we also tested larger numbers of words but found little impact on the results. The assumption here is that, in training, an association will be learnt between the occurrence of individual words and the presence of a geospatial expression or an other-spatial expression.

5.3. Word embeddings [EMBED]

While the bag of words approach represents the meaning or sense of an expression simply in terms of the words that are present, word embeddings can be regarded as introducing richer semantics, as each word in the vocabulary is represented by a multi-dimensional vector where the value of each dimension is derived from a dimensionality reduction process that learns associations between the represented word and the words with which it is commonly associated (Roweis and Saul 2000, Lavelli *et al*. 2004). The embeddings are usually learnt from very large text corpora and there are several approaches to their construction, including GloVE (Pennington *et al*. 2014), word2vec (Mikolov *et al*. 2013) and Fasttext (Bojanowski *et al*. 2017). Here, we use pre-trained GloVe embeddings. We obtain the 300-dimension embeddings of each word in an expression and average them to create a sentence-level embedding. The values of the averaged embeddings for each expression are used as the features to classify the expression. We also experimented with using the maximum of the dimensions of each word embedding rather than their average, but the results did not improve on the averaged embeddings.

5.4. Patterns [PATT]

5.4.1. Defining the patterns

We define 22 patterns to indicate the presence of types of words and groups of words, many of which we find to be more or less common in geospatial language and thus better discriminators of the class of an expression (Table 4). For example, the pattern of a preposition immediately followed by a place name is a common way to express a spatial relation, and given that patterns have been used in earlier geospatial language work (Zhang *et al*. 2009, Hall *et al*. 2011, Stock *et al*. 2013) we postulate that such patterns may be useful features for a machine learning model. Note that a few of the patterns are not specifically geospatial, as we have included patterns that are intended to assist in distinguishing geospatial from other-spatial and non-spatial expressions.

The patterns were defined manually and iteratively refined through examination of expressions in a small corpus of 500 expressions that were gathered from 39 different sources including books, news articles, photo captions (both stock photos and photo

Table 4. Summary of patterns.

Pattern	Pattern description and example
dir	Direction qualifier: *east, left and right, to and from*
gnn	Geographic noun, or place type, optionally with an adjective and a determiner: *house, the green house*
lcs	Location segment, referring to part of a geo-spatial object or place: *the western edge, the coastal side*
location	Place name: *London*
misc	MISC named entity, optionally with adjective or preposition: *On the Endeavor*
noun	Non-geospatial noun, can be qualified by an adjective: *brick, red brick*
org	Organisation: *Barclays Bank*
person	Person's name: *John Baggins*
qtf	Quantifier: *one hundred*
spatialverb	Verb that communicates location or motion: *runs*
sverb_qual_geoentity_simple	Spatial verb, optionally with adverb, followed by a geo-object type or place name: *sharply divides the mountainside*
sverb_qual_non-geoentity	As above but with non-geospatial entity
sverb_qual_rel_geoentity_complex	Spatial verb followed by qualified spatial relation term and geographic object type or place name; *sits just beside that river*
sverb_qual_rel_non-spatial-1	Spatial verb followed by a non-geospatial noun: *passed the rugby ball*
sverb_qual_rel_non-spatial-2	Spatial verb followed by two non-geospatial nouns that can be separated by a preposition: *passed the rugby ball to Harry*
sverb_to_dir	Spatial verb in combination with a directional qualifier, and optionally the *to* preposition: *travelled to the east, ran north, moved to the right*
svs_geonoun	Spatial verb satellite (in English, this is usually a preposition) followed by place type or place name: *near the house*
svs_geonoun_dir	One or more geo-object types linked by a preposition and directional qualifier: *across from the park*
svs_lcs_rel_geo-entity	Preposition followed by a location specifier and a geo-object or place name: *in the middle of the street*
svs_lcs_rel_non-geo-entity	As above but the reference object is not a geo-spatial object: *on the side of Bob*
svs_nn_geonoun	Preposition followed by spatial parthood qualifier and a place type: *in the western archipelago*
svs_noun_noun	Two nouns linked by a preposition: *tree on the mountain*

essays), and instructional/descriptive guides, and manually tagged by the authors. This corpus was entirely separate from the test data sets described in Sections 4.1–4.4. To create this corpus, we selected sources through purposive, maximum variation sampling (Patton 1990) with a view to identifying a wide range of different types of language sources. We then searched within each source manually for the first encountered instance of a geospatial sentence and then added to the corpus both that sentence and 3–4 sentences preceding and succeeding the geospatial sentence. The resulting corpus had an approximately equal division of geospatial, other-spatial and non-spatial expressions.

We built the patterns by first manually studying the expressions in this smaller corpus to create a draft set of patterns, and then progressively refined the patterns by repeatedly running the classification, studying expressions that were incorrectly classified and improving the patterns accordingly. While this approach to creation of the patterns is largely manual and there may be other, different patterns that could be defined using a different method or a different corpus, we consider that the patterns created using this method were sufficient to give an indication of the potential of a language pattern-based classification approach compared to other methods. We also consider this approach a useful addition to our experiments because the use of language patterns, and of bag of words classification methods, enables us to understand more about the nature and combinations of types of words that we see in geospatial language, in contrast to the more black-box approach of word embeddings, and some analysis of this is included in

Section 6. As demonstrated in the results section, when the predictions of pattern-based classifiers were combined into (ensemble) meta-classifiers with those of the bag of words and embedding classifiers, we obtained our best expression classification results. Our analysis of the predictive power of the individual patterns also led to their selective use as features in the creation of a metaclassifier that also provided very competitive results.

Following matching of an expression to a set of appropriate patterns, we create a matrix of expressions *vs* patterns, populated with the frequency of occurrence of a pattern within a given expression, to which we then apply a classifier. Note that other values representing the presence of a pattern were tested for use in the matrix, including presence/absence (0/1) and tf-idf, but in tuning tests, these provided no improvement.

5.5. Hybrid [EMBED-LOC-GNN-SVM]

This method is a hybrid in that it combines the embedding features with the combination of two other features representing the number of place names and the number of object type terms in the expression (which are equivalent, respectively, to the *location* and *gnn* features in the Patterns method). These features are input to an SVM classifier.

5.6. Metaclassifier [META]

The final method is a metaclassifier that is a two-stage process. In the first stage, predictions (in the form of probabilities or predicted classes) are generated from some combination of the previously described classifiers (Sections 5.1–5.4). In the second stage, these predictions are used as features by themselves or in combination with features employed in the previously described methods, as inputs either to an SVM classifier or to a simple voting system that selects the majority decision from other classifiers. It may be noted that this approach has elements of ensemble learning, in that the metaclassifiers combine predictions from several classifiers. However, as indicated, features input to the second stage can be combinations of predictions from other classifiers and of original feature data items (such as the patterns or word embeddings). We experimented with several feature combinations of classifier outputs and data items and report here on the results of three experiments that provided the best values for either precision, recall or F1 or a good overall balance.

In the first meta-classifier META-1, the inputs to an SVM classifier are the output probabilities of the three classes (geospatial, other-spatial and non-spatial) from each of the three bags of words classifiers that used Naïve Bayes, a Bayesian Network and SVM, respectively (giving 9 prediction probabilities), and the features used by the PN method (i.e. a single Boolean feature representing the presence or absence of a place name), the PATT method (the number of instances in the expression of each of 22 pattern values) and the embeddings (EMBED) method (300 values for an averaged embedding of the words in the expression).

The second meta-classifier adopts a voting system that selects the modes of predictions for each of the three output classes resulting from each of the three classifier methods of BoW-SVM, PATT-NB and EMBED-SVM. These three methods were selected as each of them provided the best average result across all classes for their respective method.

Table 5. Summary of features included for each method.

Method	Place name presence	Tf-idf word vector	Number of syntactic patterns	Averaged word embeddings	Number of place names from patterns	Number of feature types from patterns	Predicted probabilities	Predicted classes
PN	X							
BoW		X						
PATT			X					
EMBED				X				
EMBED-LOC-GNN				X	X	X		
META-1	X		X	X			BoW-NB BoW-BN BoW-SVM	
META-2								BoW-SVM PATT-NB EMBED-SVM
META-3								BoW-SVM PATT-SVM EMBED-SVM

The third meta-classifier also uses voting and is the mode of the class output predictions for each of the three classes from BoW-SVM, PATT-SVM and EMBED-SVM. Other meta-classifier approaches that we tried, but which were not better than those reported here, included using features consisting of output probabilities of all versions of BoW, PATT and EMBED; output probabilities of just the SVM versions of BoW, PATT and EMBED; probabilities of EMBED-SVM and features of BoW and PATT and the features of EMBED and PATT. A summary of which features are used in each classifier method is provided in Table 5.

6. Evaluation

6.1. Classifier performance

For each of the methods, we created a ternary classifier for the three output classes of geospatial, other-spatial and non-spatial. For all of the methods except the metaclassifiers, we ran the experiments using Naïve Bayes, Bayesian Network and SVM classifiers. These classifiers were selected as they are well known and robust and were not outperformed by a number of other classifiers that we tested. Ten-fold cross-validation was used for all methods. In the case of the baseline of the PN (Boolean place name) experiment, we only report the SVM results, as none of the other classifiers produced any better results for precision, recall or F1. For the BoW, PATT and EMBED methods we report results for all three as there was some variation in which produced the best outcomes. The results of the experiments can be seen in Table 6. A *t*-test showed that a difference in precision, recall and f-measure of less than 0.01 (1%) is not statistically significant at the 95% confidence level, so we only show figures to this level (2 decimal places). Thus, all differences between figures for different methods that can be seen in Table 6 are statistically significant.

As can be seen in Table 6, for geospatial classification, all three metaclassifiers outperform all other methods, while the hybrid classifier (EMBED-LOC-GNN-SVM) that combines features consisting of the embeddings, the numbers of place names (as represented by

Table 6. Experimental results.

	Output classes from ternary classifier									
	geospatial			other-spatial			non-spatial			
Method	p	r	f1	p	r	f1	p	r	f1	Mean
PN-B-SVM: Boolean presence of place name	.88	.65	.75	.00	.00	.00	.51	.92	.65	.48
BoW-NB: Bag of Words (BoW) Naïve Bayes	.88	.47	.61	.31	**.96**	.47	.70	.22	.34	.55
BoW-BN: Bag of Words (BoW) Bayes Net	.85	.75	.80	.76	.68	.72	.71	.84	.77	.76
BoW-SVM: Bag of Words (BoW) SVM	.83	.79	.81	.82	.78	.80	.77	.83	.80	.80
PATT-NB: Patterns Naïve Bayes	.86	.70	.77	.50	.40	.44	.61	.79	.69	.64
PATT-BN: Patterns Bayes Net	.78	.82	.80	.54	.34	.41	.65	.76	.70	.64
PATT-SVM: Patterns SVM	.76	.82	.79	.57	.26	.36	.65	.79	.71	.63
EMBED-NB: Embeddings, Glove Average, Naïve Bayes	.75	.73	.74	.56	.86	.68	.77	.55	.64	.70
EMBED-BN: Embeddings, Glove Average, Bayes Net	.84	.81	.82	.75	.86	.80	.86	.82	.84	.82
EMBED-SVM: Embeddings, Glove Average, SVM	.88	.87	.87	.87	.87	.87	.87	.88	.87	.87
EMBED-LOC-GNN-SVM: embeddings + number of place names + number of feature types	.89	.88	.89	.88	.87	.87	.88	.89	.89	.88
META-1: SVM[probabilities(BoW-NB,BN,SVM) + features(PN-B, PATT, EMBED)]	**.91**	**.90**	**.90**	.88	.89	**.88**	**.89**	.90	**.90**	**.89**
META-2: ModeOfClassPredictions(BoW-SVM + PATT-NB + EMBED-SVM)	**.91**	.86	.89	.88	.82	.85	.83	**.91**	.87	.87
META-3: ModeOfClassPredictions(BoW-SVM + PATT-SVM + EMBED-SVM)	.90	.89	.89	**.90**	.80	.85	.85	**.91**	.87	.87

the pattern *location*, see Table 4) and of geo-object types (as represented by the pattern *gnn*, see Table 4) gave the best results for the single-stage classifiers. Of the metaclassifiers, META-1 that combines BoW (bag of words) predictions with features of embeddings, patterns and place name presence-absence provides the best results for geospatial classification, with 0.91 for precision and 0.90 for recall and F1. META-2, which outputs the mode of each of the best versions of the BoW, patterns and embeddings classifiers and also provides a precision of 0.91 but with F1 of 0.89, while META-3, which takes the mode of the SVM versions of the BoW, patterns and embeddings classifiers, gives a good balance of precision (0.90) and recall (0.89) for the geospatial class. META-1 is notable for obtaining the best F1 values for all three output classes, as well as the best precision for the non-spatial class. For the other-spatial class, the best precision was obtained with META-3.

Of the individual methods of the single-stage classifiers, it is clear that the embeddings are the most successful, with the SVM version giving 0.88 for geospatial class precision and 0.87 for recall and F1. It also obtains the best F1 for both other-spatial and non-spatial classes among the individual method classifiers. A small increase in performance of the geospatial classification result for embeddings is obtained with the hybrid single-stage classifier (EMBED-LOC-GNN-SVM) that combines embeddings with features representing the numbers of place names and numbers of geographic object types in an expression, giving 0.89 for geospatial precision and F1. Among the pattern classifiers, Bayes Net provides the best geospatial F1 of 0.80. The best precision with patterns of 0.86 was obtained with the Naïve Bayes version, but with a lower recall of 0.70. Notably, the SVM version of the bag of words classifiers provides better geospatial F1 performance than any of the Pattern classifiers with an F1 of 0.81. Bags of words were, however, generally outperformed by the embeddings method, with the exception of the geospatial precision result of 0.88 with the Naïve Bayes version of bag of words. This relatively high precision was, however, accompanied by a very poor recall of 0.47. Finally, it may be noted that the simple PN baseline of the presence or absence of a place name obtained a geospatial precision of 0.88, with an F1 of 0.75. Unsurprisingly, this classifier, with only the evidence of the place name presence, was unable to classify successfully any of the other-spatial expressions. Its lower geospatial recall of 0.65 reflects the fact that it cannot identify geospatial class expressions that do not include place names and, with regard to its precision, it could misclassify an expression that mentioned a place name that was not in the context of a relative description of location.

Examination of the actual predictions that result from the different methods highlights the inability of the baseline method (PN) to identify expressions that are geospatial by our definition, but that do not include place names. For example, the following expressions were detected by META-1, but not PN:

- *A brick wall is reached beyond a bridge, with a number of obstructions over the next stretch.*
- *Thus, at first transport could follow routes used for millennia by the country's indigenous inhabitants, who had always traded small and valuable goods by foot.*

As discussed in Section 2.1, the first of these is classed as geospatial due to its mention of geographic objects (walls and bridges) with spatial relation terms. This expression is part of a textual description of the route along a canal, including place names, and, once this particular expression has been identified as geospatial, coreference resolution methods could then be used to connect this expression to the relevant place names and identify the specific geographic location described (the location of the brick wall and the bridge).

The second example would also be linked to a general geographic area through coreference resolution, but then refers to a class of objects (routes), which we consider useful as it may be combined with a spatial query to identify groups of actual geographic places.

Furthermore, the following expressions, which were classified correctly as neither geospatial nor spatial by our META-1 method, but incorrectly as geospatial by the PN method, show that our method is able to discriminate between expressions that include place names but not spatial relation terms:

- *These changes, and the need for Britain to conform with EC legislation, gave rise to ever more comprehensive labelling regulations.*
- *The US deployed the assassination as a ` despicable act of terrorism against a man of peace.*

The mention of place names in these examples occurs without any spatial relation term and thus does not meet our definition of geospatial. We exclude expressions that include only place names because methods and tools for detecting simply place names are already well-developed (Won *et al*. 2018) (see Section 2.1).

Examining the expressions that our best-performing method overall (META-1) was unable to successfully classify (which were only 9% of the total), we see that 68% of the geospatial expressions that were not detected by META-1 were classified as non-spatial and 32% as other-spatial. Several of the geospatial expressions that were misclassified included elements such as the preposition *to* (e.g. *A pilgrimage to Grantchester* *proved a literary disappointment*), most likely because while *to* can be used as a spatial preposition, it is frequently used for non-spatial purposes and thus would be less likely to be identified as a common marker of geospatial expressions by the model. Furthermore, *to* has its own part of speech tag (TO), in contrast to other prepositions (which all have the tag IN) in the Penn treebank tagging scheme[7] most commonly used by the majority of part of speech taggers, and although our PATT method (Section 5.4) takes this into account, some rules are restricted to the IN tag to avoid the loss of performance caused by the many non-spatial uses of *to*. Other types of expressions that were missed included those referring to place names that were tourist attractions rather than simply place names and thus less likely to be recognised as a place name by named entity recognition tools (e.g. *Horse Trainer Stephen Farley with Sincero at Wyong Race Club – 27 September 2011*) or implied geographic features (e.g. *After the quick sprint from Penumbra's, the cut there was bleeding heavily*). Expressions that were incorrectly classified as geospatial by the META-1 method, when they were in fact non-spatial, were mainly those that included place names with terms that could describe spatial relations, but that were instead used to describe time or attributes (e.g. *The Hop Pole is a good real ale pub with a limited range of food in the evenings*).

6.2. Feature examination

In an effort to understand which aspects of the geospatial expressions are most informative for classification, we examined the ranking of features used in the pattern classifier using all three versions of Naïve Bayes, Bayesian Network and SVM. The consistently top ranked feature was found to be the *svs_geonoun* pattern that reports the number (within an expression) of prepositional phrases consisting of a preposition followed by a place name or a geographic feature type. This is an expected outcome in that geospatial expressions in our collection are required to include a spatial relation term, and prepositional phrases are a very frequently used method for expressing spatial relations (Herskovits 1987). The top six features were the same for all three versions of the classifier (SVM, Naïve Bayes and Bayes Net) and also included the patterns corresponding to the number of occurrences of a place name, a geospatial object type, a prepositional phrase that consists of a spatial preposition followed by a directional qualifier and a place name and a verb that communicates location or motion.

When the Patterns and the Boolean place name presence features are compared alongside the embeddings and the predictions from the three versions of the bag of words classifier in the META-1 metaclassifier, five of the top six features are predictions from the bag of words classifier (see Figure 2). The second ranked feature is the prepositional phrase pattern and the seventh ranked feature is the Boolean presence of a place name, which can be regarded as intuitively reasonable in that place names are very common in geospatial expressions but are not by themselves sufficient for our definition of geospatial. The next three in the top ten ranking are an embedding feature, the prepositional phrase pattern including parthood qualifier (at number 9 in the ranking) and the number of place names.

We also performed an analysis of information gain of the individual words that make the greatest contribution to classification in the bag of words model.[8] We manually classify high information gain words by word type, as shown in Table 7, and display the

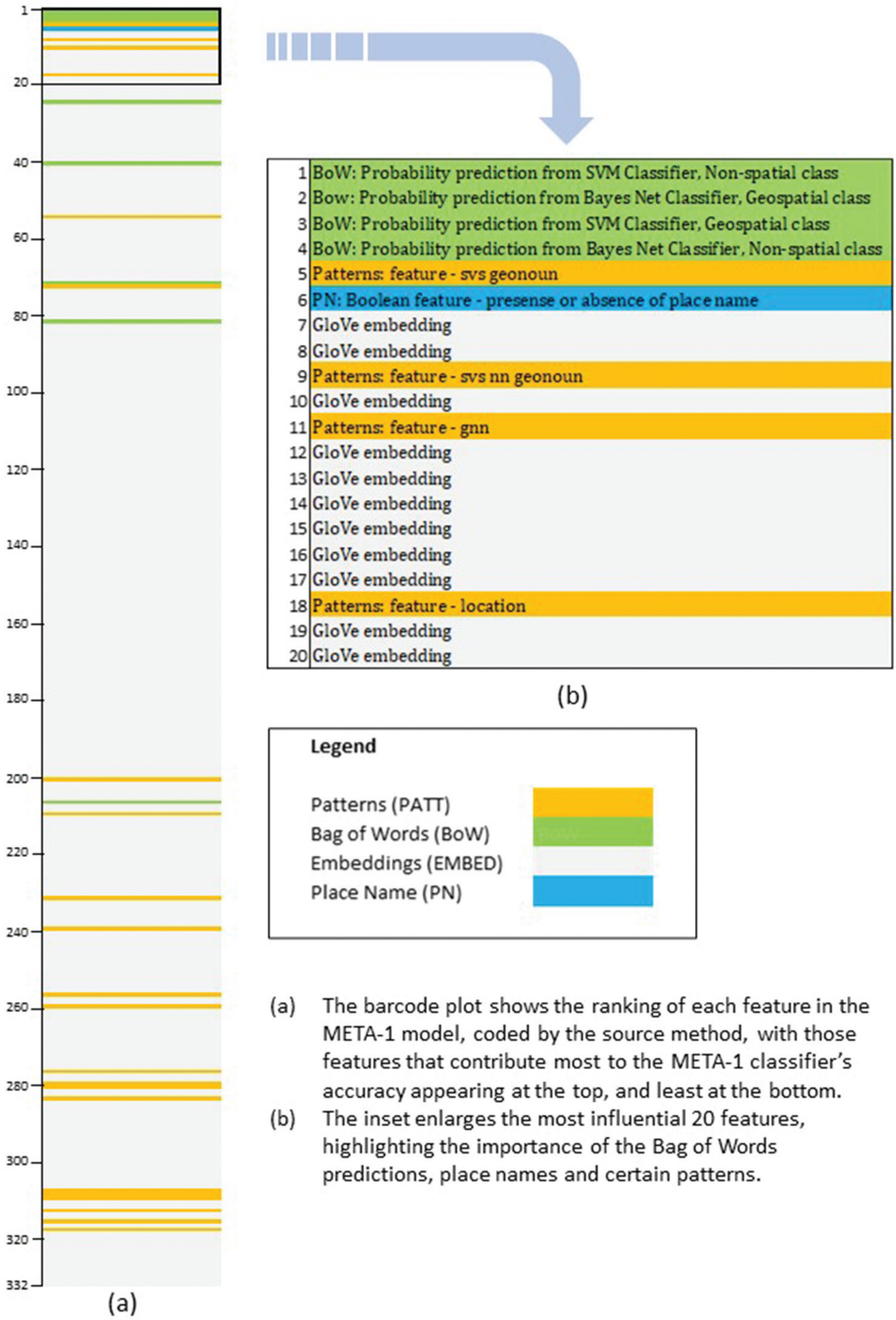

Figure 2. Feature ranking in META-1 model.

proportions of different word groups in Figure 3. The classification scheme is subjective, based on examination of the words that appear in the lists and their groupings, and we analyse the 60 words that provide the most information gain (listed top to bottom, left to right in Table 7). This analysis shows some clear patterns of the kinds of words that are important contributors to particular classes in the bag of words approach. As seen in Table 7 by counts and also graphed in Figure 3, for the geospatial class, geographic object types are most important, while spatial relation terms and other locative words are surprisingly few in the top 60 list, in contrast to the other-spatial class, which relies much more heavily on spatial relation terms and other locative words, alongside body parts and indoor objects, as would be expected. The highly ranked other-spatial words do include a few geospatial terms, which might reflect the fact that our corpus is biased towards expressions with geographic content, but not all of these are locative with respect to geographic features. There are less clear distinctions for the non-spatial class, in which several groups of words are important, with no dominant group. The presence of geospatial and other-spatial terms here can be attributed to the use of these words in discriminating the other-spatial and geo-spatial expressions from the non-spatial expressions. It may be noted that while many of the words that help discriminate between the classes here are generic, some of them can be regarded as distinctive to our dataset (as, for example, in the case of particular place names).

We can see the role of the patterns and distinctive words in the correct classification by META1 of geospatial example expressions 3 and 4 in Section 2.1, neither of which contain explicit place names. Thus, *Climb the hill and enter the echoing fifteenth-century Gothic church to peer through glass panels at the medieval foundations* was detected as containing the *svs-geonoun* pattern (*at* and *foundations* in *at the medieval foundations*) along with several of the pattern *gnn* with a geo-feature type (*hill, church* and *foundations*). Notably, *hill* and *church* were also highly ranked in the bag of words methods (Table 7) and can be expected therefore to have been ranked highly in their associated word embeddings. Similarly, in the expression *This fort is in the centre of the city near the sea-front and in front of it is a very pleasant tree-lined road with an open-air café,* the *svs-geonoun* pattern is again present,in this case, three times, including *near the sea-front*, while six *gnn* was detected (*fort, city, sea, tree*, *road* and*café*). The other two geospatial expression examples (*Maybe, thousands of years* . . . . and *About 100 homes across* . . . *.)* were also correctly classified by the META1 classifier, as they were by the baseline as they contain place names.

All three other-spatial expressions in Section 2.2 (*Masklin darted between . . . ; Reaching for a napkin . . . ; Frank stood up . . .)* were correctly classified by the META1 classifier. While all have spatial relation terms, none of them contains the *svs-geonoun* pattern or the patterns of geospatial feature types (*gnn*) or *location*. Notably, all three have at least one of the *person* patterns and at least one of the top ranked other-spatial bag of words terms (*seat*, threw, beneath and *hand*). The second and third of these examples also contain other body parts that can be expected to be captured in meaning by their respective word embeddings. Regarding the two non-spatial expressions from Section 2.3 (*He's even been stealing from the business to get cash to buy his drugs and You look as if you've been bickering with a volcano)*, both were correctly classified by the META1 classifier. Neither has typical spatial relation terms, with the possible exception of *with*, as reflected in the

Table 7. Words that provide most information gain in the bag of words method.

	geospatial			other-spatial			non-spatial		
	road	miles	`	door	sitting	city	road	london	miles
	south	st	sydney	room	neck	fire	south	floor	back
	park	beach	tel	hand	walls	street	park	valley	nottingham
	river	topics	path	table	threw	town	river	train	table
	north	area	map	back	river	turned	door	centre	forest
	street	mountain	canal	floor	stood	fingers	street	wall	hand
	east	church	door	eyes	*	topics	north	water	square
	city	forest	africa	round	flung	-	east	social	map
	station	nottingham	county	road	put	stairs	side	mountain	location
	-	hill	trees	sat	national	picked	station	beach	&
	london	highway	house	feet	chair	desk	city	window	trees
	west	sea	bus	hair	east	mouth	`	car	small
	village	europe	left	hands	london	&	room	st	bridge
	town	location	built	window	north	top	village	town	island
	valley	trail	australia	bed	bathroom	seat	area	feet	bed
	lake	national	land	south	waist	holding	-	sat	turned
	train	country	located	park	arm	west	west	terms	tel
	walk	bridge	march	face	corridor	pocket	walk	topics	highway
	centre	railway	room	head	kitchen	arms	round	located	hall
	&	island	mountains	beneath	looked	tears	lake	railway	sea
Geographic features and place names	38			7			25		
Directions	4			4			4		
Spatial relations and other locative words	6			12			9		
Body parts	0			13			2		
Indoor features	2			14			8		

detected patterns in which the first has only the pattern of *noun* (of which there are two), while the second has just the pattern of one *gnn* (*volcano*). Neither of them contains terms that are highly ranked in the bag of words method (Table 7) as either geospatial or other-spatial.

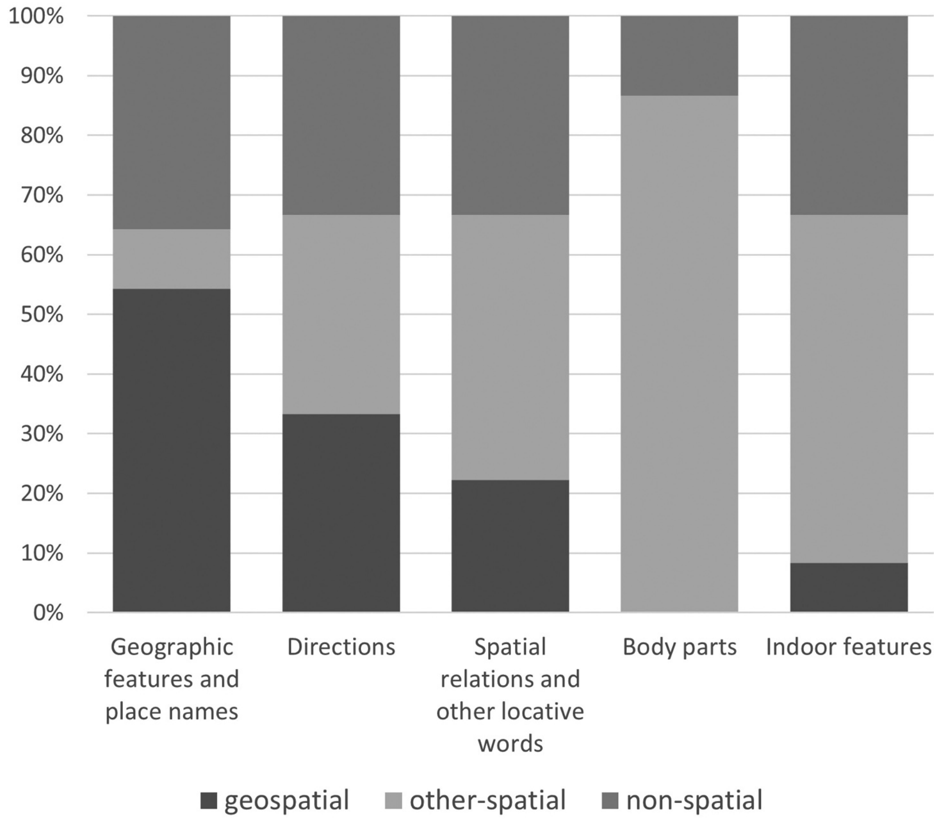

Figure 3. Proportions of word types in each class.

6.3. Testing with additional data sets

We further tested the model trained on the Mechanical Turk data set against the two additional data sets described in Section 3.3 (the NIWA data set) and 3.4 (the COVID data set). The results for the META-1 method (overall, the best performing method for the MT6.5K data set), are shown in Table 8. Similar to the MT6.5K data set, results for the other metaclassifier and embedding methods were slightly lower and the individual methods did not perform as well as the metaclassifier and embedding methods.

Table 8. META-1 results for MT6.5K, NIWA and COVID data sets.

	Output classes from ternary classifier									
	geospatial			other-spatial			non-spatial			
Method	p	r	f1	p	r	f1	p	r	f1	Mean
Ten Fold Cross Validation with MT6.5K (from Table 5)	.91	.90	.90	.88	.89	.88	.89	.90	.90	.89
Trained on MT 6.5 K data, tested on NIWA	.88	.75	.81	.34	.66	.44	.59	.58	.58	.63
Ten Fold Cross Validation with NIWA	.89	.92	.90	.49	.40	.44	.71	.69	.70	.68
Trained on MT 6.5 K data, tested on COVID	.74	.65	.69	.50	.26	.34	.66	.77	.71	.59
Ten Fold Cross Validation with COVID	.78	.81	.79	.35	.16	.22	.76	.79	.77	.60

Table 9. Example expressions from the COVID data set.

Expression	Manual Classification	META-1 Classification (ten-fold cross validation)
All over the United States, people are fleeing urban areas with high infection rates for the perceived safety and natural beauty of rural areas	Geospatial	Geospatial
At a small backwoods inn on the banks of the Rogue River in Oregon, 200 miles from the nearest hospital, the owner told me that even with a NO VACANCY sign up, and the restaurant closed down, she still had people knocking on the door, looking for a long-term place to stay	Geospatial	Geospatial
Leaves will be picked up between April 20 and May 1, and brush bundles will be collected between May 4 and May 8 in conjunction with regular trash and recycling schedules.	Other-spatial (leaves, brush, recycling are all mobile, impermanent)	Non-spatial
Power was cut when the meter board flew to bits, and windows in the room fell in.	Other-spatial	Other-spatial
I want to tell you tonight that I don't believe for one minute that that won't be extended at least until the end of April, Vigeant told city councilors during Monday night's meeting.	Non-spatial	Non-spatial
The number of people recovered is based on the number of confirmed positive cases.	Non-spatial	Non-spatial

As can be seen, META-1 was able to predict the geospatial class for the NIWA data set almost as well as for the MT6.5K data set using ten-fold cross validation, with an F1 value of 0.90. The model that was trained on the MT6.5K data set and used to classify the NIWA data set also had high precision (0.88), but lower recall (0.75). Classification for the geospatial class was less successful for the COVID data set, although both precision and recall were still close to 0.8 (0.78 and 0.81 respectively) for ten-fold cross validation. Both were slightly lower (0.74) for the model trained on the MT6.5K data set. Results for the other-spatial class are poor across both the NIWA and COVID data sets and this is likely to be due to the small number of expressions in that class, making effective training difficult.

It is clear from these results that our methods were less successful at classifying the COVID data set than the NIWA or MT6.5K data sets. The COVID data set was scraped from the web and, in addition to text in standard prose form (including the examples shown in Table 9), included content such as table headings, lists of place names and figures (case numbers), special characters and website navigation instructions. As mentioned in Section 3.5, there were also challenges with metonymic language, in which a place

Table 10. Example expressions from the NIWA data set.

Expression	Manual Classification	META-1 Classification (ten-fold cross validation)
About 100 homes across Te Puke, Omanu, Matapihi, and Paengaroa were without power on the night of the 19th.	Geospatial	Geospatial
Snow was lying in bush above Marsden Valley on the Barnicoat Range at a height of 650 m at 6am on the 18th.	Geospatial	Geospatial
Power lines had been ripped out, windows smashed, and entire roofs had blown away.	Other-spatial	Geospatial
The strong winds sent calves flying into the air, and horses were hurt when they were struck by lightning.	Other-spatial	Non-spatial
Other farmers had to hire generators until power was restored late on the 19th.	Non-spatial	Non-spatial
Power cuts and electrical faults were experienced.	Non-spatial	Non-spatial

name was used to represent an organisation, rather than a physical entity, and such cases are a challenge for automated methods to detect. Table 9 shows the manual classification as well as the classification of our META-1 method. Most of these examples were correctly classified, but one of the other-spatial expressions was misclassified as non-spatial. This might have been in part due to the use of less common spatial relation terms *picked up* and *collected*.

In contrast, the NIWA data set (see examples shown in Table 10) contained text extracted from reports that were intended to describe impacts in textual form, and the content was thus cleaner. The challenge with the NIWA data set was that it was heavily skewed towards geospatial expressions, with much lower numbers of other-spatial and non-spatial expressions, and it contained frequent references to weather, which were scarce in the MT6.5K data set (explaining lower recall when the model trained on the MT6.5K data set was used to classify the NIWA data set). Weather expressions are challenging because they do refer geographic objects, but often in combination with a mobile weather event (wind, storm and rain), which by our definition is not defined as geospatial. Table 10 illustrates the challenges of other-spatial expressions in particular. The first other-spatial example refers to power lines and roofs, which are commonly associated with geospatial expressions (although in this case, the reference objects of windows and roofs do not qualify as geospatial by our definition). The second other-spatial example uses less-common spatial relation terms *sent flying* and *struck by*, which present a challenge for the automated classifier.

6.4. Distinguishing geospatial expressions according to requirement for coreference

The results of the META-1 classifier have demonstrated its power to identify geospatial expressions that include spatial relation terms but do not contain place names. As indicated previously, such expressions are of interest as they have the potential to be georeferenced, provided that the reference objects of spatial relation terms could be resolved, through a process of coreference, to a named place for which coordinates are available. Given that it might be of interest to distinguish explicitly the expressions having the potential for direct georeferencing, as opposed to those that would require coreference, it can be noted that our PN classifier provides a means to do this. Thus, once an expression has been classified as geospatial by the META-1 classifier, it could be further sub-classified as subject to direct georeferencing, if it was also classified as geospatial by the PN classifier (which is based on identifying whether an expression contains a place name).

7. Conclusions

In this paper, we have presented methods for detecting the presence of geospatial expressions that contain relative descriptions of the locations of geographical phenomena. We describe the process of creating a test collection of geospatial, other-spatial and non-spatial expressions and we then use that collection, along with two other data sets, to experiment with the development of automated methods for detecting and distinguishing between locative expressions. This work is motivated by the need (not addressed in this paper) to develop methods for georeferencing texts in which locations are described relatively through the use of spatial relation terms. Given the low incidence of such geospatial expressions in generic text, our methods serve the purpose of selecting candidate expressions for georeferencing and hence filtering out the large number of expressions that are not candidates for georeferencing. A subsequent parsing step can then be applied to identify and determine the coordinates of place names using named entity recognition and toponym resolution methods (Karimzadeh *et al*. 2019) and to resolve coreferences where required (Manzoor and Kordjamshidi 2018, Joshi *et al*. 2019, Sukthanker *et al*. 2020) (e.g. for expressions that do not include a place name in the sentence itself, but elsewhere in the text), before the application of models for georeferencing phrases that include spatial relation terms (Doherty *et al*. 2011, Hall *et al*. 2011, Wanichayapong *et al*. 2011, Bahir and Peled 2016, Chen *et al*. 2018). These steps are the subject of ongoing research.

As part of the process of creating our test collection, we have presented a simple definition of a geospatial expression that requires the reference object to be geographic but that does not otherwise constrain the way in which spatial relationships are communicated. We have used this definition, along with complementary definitions of other-spatial expressions and non-spatial expressions, as part of the instructions for a web-based annotation exercise using Mechanical Turk. To ensure consistency in the final data set, we applied a procedure to eliminate annotators who were very inconsistent and to remove expressions that were subject to considerable disagreement in their annotation. Clearly, ambiguous expressions could be of interest in their own right as part of a corpus of different types of locative expressions, but their removal here was motivated by the requirement to produce a dataset that could be used effectively for automated classification.

For the purpose of automated detection of geospatial expressions and their distinction from other-spatial and non-spatial expressions, we experimented with several types of classifiers that employed various features derived from the natural language expressions. These classifiers can be grouped into single-stage classifiers that use features derived from the expressions and two-stage metaclassifiers that use features as either combinations of class predictions from the single-stage classifiers or combinations of class probabilities in combination with features from the single-stage classifiers. The single-stage features consisted of the average word embeddings of all words in an expression, a bag of words vector that records the *tf-idf* values of those words in the entire document collection that are present in an individual expression, a set of values representing the presence of language patterns designed to detect particular characteristics of geospatial language and a Boolean feature indicating the presence or absence of place names. The best performing method for identifying geospatial expressions, with 0.91 for precision and 0.90 for both recall and F1 values, used a two-stage meta-classifier that combined output class probability predictions from bag of words classifiers with other features consisting of averaged embeddings, the set of patterns and the Boolean feature, indicating the presence or absence of place names. This classifier provided the best F1 values for each of the geospatial, other-spatial and non-spatial classes (0.90, 0.88 and 0.90, respectively), and the best precision for both geospatial (0.91) and non-spatial (0.89) classes. The top precision, recall and F1 values for all three classes were all achieved by our two-stage meta-classifiers, with one exception: the best recall for the other-spatial class (0.96) was achieved by a Naïve Bayes bag of words classifier, but this classifier performs poorly in precision, while the meta-classifiers provide a better balance across precision and recall.

Analysis of the patterns that ranked most highly across different forms of classifiers found that, purely among the patterns, the prepositional phrase was the highest ranked, while other patterns representing place names, geospatial object types and parthood qualified prepositional phrases were consistently in the top 5 features. In our best classifier (META-1), five of the top six features were bag of words predictions and the other was the prepositional phrase pattern, while the seventh was the Boolean place name. For example, the geospatial expression *Climb the hill and enter the echoing fifteenth-century Gothic church to peer through glass panels at the medieval foundations* contains the *svs-geonoun* pattern (*at the medieval foundations*) and several *gnn* geo-feature types (*hill, church* and *foundations*) of which *hill* and *church* were high ranking in the bag of words methods. In contrast, the other-spatial expression *Masklin darted between two seats, around a pair of giant shoes, and threw himself flat on the carpet,* while having a spatial relation term, does not contain the *svs-geonoun* pattern or *gnn* geospatial feature type patterns. It also contains two of the top ranked other-spatial bag of words terms (*seat* and *threw*) and other non-geographical objects (*shoes* and *carpet*). Analysis of the words ranked most highly in the bag of words method revealed that, for geospatial classification, they were geospatial objects and place names, while for other-spatial classification, it was words that relate to indoor phenomena, in combination with spatial relation terms, that were most prevalent.

The methods described in this paper may be applied to identify geospatial language from a range of text sources including social media, blogs, environmental reports, newspaper articles from the web, text archives and other sources, and we have demonstrated this with two additional data sets from the domains of weather and health (COVID-19). It is likely that results on other, more specialised document types like microblogs and short-form social media (e.g. Twitter) could be further improved with additional training data.

Our methods go beyond previous work in that they detect not just prepositions but also a wider range of syntactic forms of geospatial language; they distinguish geospatial from other-spatial language. Our methods complete a first step that can then be followed by the applications of methods for georeferencing more complex text location expressions.

Notes

1. Note that we do not consider spatial language at scales more extensive than geographic, such as at astronomic scales or at finer levels than indoor or table-top.
2. http://www.clres.com/prepositions.html
3. http://geospatiallanguage.massey.ac.nz/
4. https://www.mturk.com/
5. https://hwe.niwa.co.nz/
6. https://bootcat.dipintra.it/
7. https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html
8. Note that while embeddings provided better performance than Bag of Words their individual dimensions are not interpretable in terms of the vocabulary of the expressions
9. This class is referred to as other-spatial in the paper.
10. This class is referred to as non-spatial in the paper.

Contributions

Kristin Stock conducted most of the machine learning and feature examination, prepared and annotated data sets and drafted the paper.

Christopher Jones drafted the paper, annotated data sets and advised on methods.

Shaun Russell designed and developed the patterns used in one of the methods and contributing to the meta-methods.

Mansi Radke annotated data sets and developed methods to extract some baseline features of the model.

Prarthana Das annotated data sets and developed methods to extract some baseline features of the model.

Niloofar Aflaki extracted embeddings for the data sets.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Kristin Stock is Director of the Massey Geoinformatics Collaboratory and a Senior Lecturer in Information Technology at Massey University. She completed a surveying degree at the University of Otago, graduate diplomas in urban and regional planning and surveying practice, and a PhD in semantic database integration from Queensland University of Technology. She has also worked in the private sector and government, designing a Land Information System for the Australian Capital Territory. Her research interests are in geospatial natural language, semantics and data integration.

Christopher B. Jones is Professor of Geographical Information Systems at Cardiff University, having worked previously at the University of South Wales, the University of Cambridge, BP Exploration and the British Geological Survey. He studied geology at Imperial College and Bristol University and has a PhD on periodicities in fossil growth rhythms from the University of Newcastle upon Tyne. Current research focuses on geographical information retrieval and earlier research on cartography led to the development of the Maplex cartographic name placement software.

Shaun Russell has completed a Bachelor of Information Sciences at Massey University. He is currently working as a Consultant at a leading Legal Software firm, specialising in Data Analytics, Reporting, and Client Customization Services.

Mansi Radke is an assistant professor in the Department of Computer Science and Engineering at Visvesvaraya National Institute of Technology (VNIT), Nagpur, Maharashtra, India. She has completed her Masters from University of Maryland, Baltimore County and Ph.D. from VNIT. She has industry experience of 5+ years too. Her interests are in the areas of Information Retrieval and Natural language processing.

Prarthana Das has completed masters in Computer Science and Engineering from Visvesvaraya National Institute of Technology (VNIT), Nagpur. She has 3 years experience in the IT industry and is currently working as an Associate Applications Developer at Oracle Financial Services Software Private Limited specializing on iOS App Development.

Niloofar Aflaki is a Ph.D. candidate in IT. Her Ph.D. research focuses on developing new methods to interpret natural language location descriptions. Her interests are natural language processing, machine learning, deep learning, and programming, especially with python.

ORCID

Christopher B. Jones http://orcid.org/0000-0001-6847-7575

Data and codes availability statement

The data and code that support the findings of this study are available at https://doi.org/10.6084/m9.figshare.12561983

Ethical statement

Ethical approval for the research reported in this paper was obtained from Massey University (approval number 4000023606). The research was classed as low-risk.

Aflaki, N., Russell, S., and Stock, K., 2018. Challenges in creating an annotated set of geospatial natural language descriptions (Short Paper). *In*: S. Winter, A. Griffin, and M. Sester, eds. *10th International Conference on Geographic Information Science (GIScience 2018)*. Vol. 114. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 20: 1–20:6. doi:10.4230/LIPIcs.GISCIENCE.2018.20.

Bahir, E. and Peled, A., 2016. Geospatial extreme event establishing using social network's text analytics. *GeoJournal*, 81 (3), 337–350. doi:10.1007/s10708-015-9622-x.

Bojanowski, P., *et al*., 2017. Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5, 135–146. doi:10.1162/tacl_a_00051

Carlson-Radvansky, L.A., Covey, E.S., and Lattanzi, K.M., 1999. "What" effects on "Where": functional influences on spatial relations. *Psychological Science*, 10 (6), 516–521. doi:10.1111/1467-9280.00198.

Chen, H., Winter, S., and Vasardani, M., 2018. Georeferencing places from collective human descriptions using place graphs. *Journal of Spatial Information Science*, 2018 (17), 31–62. doi:10.5311/JOSIS.2018.17.417.

Clematide, S., *et al*. 2018. *Crowdsourcing toponym annotation for natural features: how hard is it?* Corpus in GIScience: Going Beyond Butterfly Collecting, Workshop at GIScience. Melbourne, Australia.

Coventry, K.R. and Garrod, S.C., 2004. *Saying, seeing and acting: the psychological semantics of spatial prepositions*. Psychology Press. doi:10.4324/9780203641521.

Dittrich, A., *et al*. 2015. A Classification Schema for Fast Disambiguation of Spatial Prepositions. *In*: *Proceedings of the 6th ACM SIGSPATIAL International Workshop on GeoStreaming*. Seattle, US. 78–86. doi:10.1145/2833165.2833167.

Doherty, P., *et al*., 2011. Georeferencing incidents from locality descriptions and its applications: a case study from Yosemite National Park Search and Rescue. *Transactions in GIS*, 15 (6), 775–793. doi:10.1111/j.1467-9671.2011.01290.x.

Garner, B., 2017. Grammar and usage. *In*: *The Chicago Manual of style online*. Chicago: University of Chicago Press. Available from: https://www.chicagomanualofstyle.org .

Gritta, M., *et al*., 2018. What's missing in geographical parsing? *Language Resources and Evaluation*, 52 (2), 603–623. doi:10.1007/s10579-017-9385-8.

Hall, M., Smart, P.D., and Jones, C.B., 2011. Interpreting spatial language in image captions. *Cognitive Processing*, 12 (1), 67–94. doi:10.1007/s10339-010-0385-5.

Hassani, K. and Lee, W.-S. 2017. Disambiguating spatial prepositions using deep convolutional networks. *In*: *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, San Francisco, US, 3209–3215.

Herskovits, A., 1987. *Language and Spatial Cognition*. Cambridge, UK: Cambridge University Press.

Hudelot, C., Atif, J., and Bloch, I., 2008. Fuzzy spatial relation ontology for image interpretation. *Fuzzy Sets and Systems*, 159 (15), 1929–1951. doi:10.1016/j.fss.2008.02.011.

Jackendoff, R., 1983. *Semantics and cognition*. Cambridge, Massachusetts: The MIT Press.

Joshi, M., *et al*. 2019. BERT for Coreference Resolution: baselines and Analysis. *In*: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP*), Hong Kong, China. 5803–5808. doi:10.18653/v1/D19-1588.

Karimzadeh, M., *et al*., 2019. GeoTxt: a scalable geoparsing system for unstructured text geolocation. *Transactions in GIS*, 23 (1), 118–136. doi:10.1111/tgis.12510.

Khan, A., Vasardani, M., and Winter, S. 2013. Extracting spatial information from place descriptions. *COMP'13 Proceedings of The First ACM SIGSPATIAL International Workshop on Computational Models of Place*, Orlando, Florida, US, 62–69.

Kordjamshidi, P., van Otterlo, M., and Moens, M.-F., 2017. Spatial role labeling annotation scheme. *In*: N. Ide and J. Pustejovsky, eds. *Handbook of linguistic annotation*. Netherlands: Springer, 1025–1052. doi:10.1007/978-94-024-0881-2_38.

Kordjamshidi, P., Roth, D., and Moens, M.-F., 2015. Structured learning for spatial information extraction from biomedical text: bacteria biotopes. *BMC Bioinformatics*, 16 (1), 129. doi:10.1186/s12859-015-0542-z.

Kordjamshidi, P., Van Otterlo, M., and Moens, M.-F., 2011. Spatial role labeling: towards extraction of spatial relations from natural language. *ACM Transactions on Speech and Language Processing*, 8 (3), 4:1–4:36. doi:10.1145/2050104.2050105.

Kray, C., *et al*., 2013. Transitional spaces: between indoor and outdoor spaces. *In*: T. Tenbrink, *et al*., eds. *Spatial information theory*. New York: Springer International Publishing, 14–32.

Kunze, L., Doreswamy, K.K., and Hawes, N. 2014. Using Qualitative Spatial Relations for indirect object search. *In*: *2014 IEEE International Conference on Robotics and Automation (ICRA*). 163–168. doi:10.1109/ICRA.2014.6906604

Lavelli, A., Sebastiani, F., and Zanoli, R. 2004. Distributional term representations: an experimental comparison. *In*: *Proceedings of the Thirteenth ACM International Conference on Information and Knowledge Management*. Washington DC, US, 615–624. doi:10.1145/1031171.1031284.

Levinson, S.C., 2003. *Space in language and cognition: explorations in cognitive diversity*. Vol. 5. Cambridge, UK: Cambridge University Press.

Litkowski, K. and Hargraves, O. 2007. SemEval-2007 Task 06: word-sense Disambiguation of Prepositions. *In*: *Proceedings of the 4th International Workshop on Semantic Evaluations*. Prague, Czech, 24–29. Available from: http://dl.acm.org/citation.cfm?id=1621474.1621479 .

Liu, F. 2013. *Automatic identification of locative expressions from informal text*. Masters by Coursework & Shorter thesis. University of Melbourne. Available from: http://minerva-access.unimelb.edu.au/handle/11343/38520 .

Liu, F., Vasardani, M., and Baldwin, T. 2014. Automatic identification of locative expressions from social media text: a comparative analysis. *In*: *Proceedings of the 4th International Workshop on Location and the Web*. Shanghai, China, 9–16. doi:10.1145/2663713.2664426.

Luhn, H.P., 1957. A statistical approach to mechanized encoding and searching of literary information. *IBM Journal of Research and Development*, 1 (4), 309–317. doi:10.1147/rd.14.0309.

Manzoor, U. and Kordjamshidi, P. 2018. Anaphora resolution for improving spatial relation extraction from text. *In*: *Proceedings of the First International Workshop on Spatial Language Understanding*. New Orleans, US, 53–62. doi:10.18653/v1/W18-1407.

Melo, F. and Martins, B., 2017. Automated geocoding of textual documents: a survey of current approaches. *Transactions in GIS*, 21 (1), 3–38. doi:10.1111/tgis.12212.

Mikolov, T., *et al*., 2013. Distributed representations of words and phrases and their compositionality. *In*: C.J.C. Burges, *et al*., eds. *Advances in neural information processing systems 26*. Curran Associates, Inc, 3111–3119. Available from: http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf .

Montello, D.R., 1993. Scale and multiple psychologies of space. *In*: A.U. Frank and I. Campari, eds. *Spatial information theory a theoretical basis for GIS*. Berlin: Springer, 312–321.

Patton, M.Q., 1990. *Qualitative evaluation and research methods*. Newbury Park, California, US: SAGE Publications, inc.

Paul, R., *et al*., 2018. Efficient grounding of abstract spatial concepts for natural language interaction with robot platforms. *The International Journal of Robotics Research*, 37 (10), 1269–1299. doi:10.1177/0278364918777627.

Pennington, J., Socher, R., and Manning, C. 2014. Glove: global Vectors for Word Representation. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP*), Doha, Qatar, 1532–1543. doi:10.3115/v1/D14-1162.

Potthast, M. 2010. Crowdsourcing a Wikipedia Vandalism Corpus. *In*: *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, Geneva, Switzerland, 789–790. doi:10.1145/1835449.1835617.

Radke, M., *et al*., 2019. Detecting the geospatialness of prepositions from natural language text (Short Paper). *In*: S. Timpf, *et al*., eds. *14th international conference on spatial information theory (COSIT 2019)*. Vol. 142. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 11: 1–11:8. doi:10.4230/LIPIcs.COSIT.2019.11.

Roweis, S.T. and Saul, L.K., 2000. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290 (5500), 2323–2326. doi:10.1126/science.290.5500.2323.

Salton, G. and Buckley, C., 1988. Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24 (5), 513–523. doi:10.1016/0306-4573(88)90021-0.

Spranger, M., Suchan, J., and Bhatt, M. 2016. Robust natural language processing: combining reasoning, cognitive semantics, and construction grammar for spatial language. *In*: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, New York, US, 2908–2914. Available from: http://dl.acm.org/citation.cfm?id=3060832.3061028 .

Stock, K., *et al*., 2013. Creating a corpus of geospatial natural language. *COSIT 2013: Proceedings of the Conference on Spatial Information Theory.*Scarborough, UK. Lecture Notes in Computer Science (LNCS) 8116, 279–298. doi:10.1007/978-3-319-01790-7_16.

Sukthanker, R., *et al*., 2020. Anaphora and coreference resolution: a review. *Information Fusion*, 59, 139–162. doi:10.1016/j.inffus.2020.01.010

Surdeanu, M., *et al*. 2012. Multi-instance multi-label learning for relation extraction. *In*: *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. Jeju Island, Korea, 455–465. Available from: https://www.aclweb.org/anthology/D12-1042 .

Talmy, L., 1983. How language structures space. *In*: H.L. Pick and L.P. Acredolo, eds. *Spatial orientation: theory, research, and application*. Springer US, 225–282. doi:10.1007/978-1-4615-9325-6_11.

Talmy, L., 2000. *Toward a cognitive semantics*. Vol. 2. Cambridge, Massachusetts, US: MIT press.

Tenbrink, T., 2011. Reference frames of space and time in language. *Journal of Pragmatics*, 43 (3), 704–722. doi:10.1016/j.pragma.2010.06.020.

Tyler, A. and Evans, V., 2003. *The semantics of English prepositions: spatial scenes, embodied meaning, and cognition*. Cambridge, UK: Cambridge University Press.

Wallgrün, J.O., *et al*., 2014a. Construction and first analysis of a corpus for the evaluation and training of Microblog/Twitter geoparsers. *Proceedings of the 8th Workshop on Geographic Information Retrieval*, New York, US, 1–8. doi:10.1145/2675354.2675701

Wallgrün, J.O., *et al*., 2018. GeoCorpora: building a corpus to test and train microblog geoparsers. *International Journal of Geographical Information Science*, 32 (1), 1–29. doi:10.1080/13658816.2017.1368523.

Wallgrün, J.O., Klippel, A., and Baldwin, T., 2014b. Building a corpus of spatial relational expressions extracted from web documents. *Proceedings of the 8th Workshop on Geographic Information Retrieval*, 6, 1–6:8. doi:10.1145/2675354.2675702

Wang, J. and Hu, Y., 2019. Enhancing spatial and textual analysis with EUPEG: an extensible and unified platform for evaluating geoparsers. *Transactions in GIS*, 23 (6), 1393–1419. doi:10.1111/tgis.12579.

Wanichayapong, N., *et al*. 2011. Social-based traffic information extraction and classification. *In*: *2011 11th International Conference on ITS Telecommunications*. 107–112. doi:10.1109/ITST.2011.6060036.

Wolf, S.J., Henrich, A., and Blank, D., 2014. Characterization of toponym usages in texts. *Proceedings of the 8th Workshop on Geographic Information Retrieval*, New York, US, 7, 1–8. doi:10.1145/2675354.2675703

Won, M., Murrieta-Flores, P., and Martins, B., 2018. Ensemble Named Entity Recognition (NER): evaluating NER tools in the identification of place names in historical corpora. *Frontiers in Digital Humanities*, 5. doi:10.3389/fdigh.2018.00002.

Wong, B.T. and Lee, S.Y. 2013. Annotating legitimate disagreement in corpus construction. *In*: *Proceedings of the 11th Workshop on Asian Language Resources*, Nagoya, Japan, 51–57.

Zhang, C., *et al*. 2009. Rule-based extraction of spatial relations in natural language text. *In*: *CiSE 2009. International Conference on Computational Intelligence and Software Engineering*, Wuhan, China, 1–4.

Appendices

Appendix A. Mechanical Turk instructions to workers

Geospatial

Geospatial expressions have the following characteristics:

(1) They include a word or group of words (a spatial relation) that describes the location or movement of one object (the located object) relative to another (the reference object).

The word that describes the spatial relation between the located and reference objects is commonly a preposition, but may also be a verb etc.

For example, the house was **in** the country, the road **ran beside** the Thames.

(2) The reference object of the spatial relation ('the country' and 'the Thames' in the examples above) is found outdoors or in large public places, is static in nature and unlikely to move in the normal course of events AND is of a scale that is likely to occur on a map. This may range in scale from objects such as street furniture (lamp posts, fire hydrants) up to objects on a global scale. For example, the house was in the **country**, the road ran beside the **Thames**.

Spatial[9]

Like geospatial expressions, spatial expressions contain a spatial relation using the same definition as for geospatial expressions. However, in the case of spatial expressions, the reference object of the spatial relation does not meet the criteria for geospatial expressions (item 2 above).

Thus, reference objects may be indoor, mobile or small scale. For example, the book is **lying on** the **table**; the dog **sits in** the **car**.

If an expression contains both spatial and geospatial elements, it should be classed as geospatial

Neither[10]

Everything that is not geospatial or spatial should be classed as neither.

See more examples of expressions in each class (this is a hyperlink to the web page shown in Appendix B).

Appendix B. Additional examples

Geospatial Sentences

Include the following:

- a spatial relation that describes the location or movement of an object relative to another reference object.

- a geographic reference object that is found outdoors or in large public places, static and unlikely to move.

Examples:

- First-class rail travel in Portugal, Spain, Italy and Eastern Europe is particularly of good value because their rail fares are so cheap to start with
- A brick wall is reached beyond a bridge with a number of obstructions over the next stretch
- A broad gravel terrace runs parallel with the house and terminates at a small group of Scotch firs, which is immediately approached by a flight of grass steps
- The number of patients in the state hospitalized with the virus, 1,592, is 58 less than a day earlier.
- A three-car crash on the Auckland Harbour Bridge closed one northbound lane.

Spatial Sentences
Include the following:

- a spatial relation that describes the location or movement of an object relative to another reference object.
- a reference object that is indoors, mobile or small scale.

Examples:

- A nurse smiled and pulled faces at a friend just before a van ploughed into the back of her car causing fatal injuries
- A plump figure in a dress of olive green silk came towards me from the dining room walking so quickly that she skidded and slithered in her evening slippers on the marble floor
- After the battle, the golden crown was taken from the helm of the fallen Plantagenet, the last of his line and placed upon the head of the first of our Tudor kings
- The floor of the bar and pokie room was covered in several centimetres of water.

Neither Sentences
Do not include a spatial relation.
May include a place name, but without a spatial relation.

Examples:

- Now is the time to look into being a poll worker during the upcoming elections.
- Your information is strictly confidential and will be treated as the private medical record it is.
- A pair of intense fronts on the afternoon and night of the 14th.
- They had not had any contact from support agencies or civil defence services.
- 'And now if we could get back to the point, I believe you were about to explain to me about this proposition of yours.'