Spatial Cognition & Computation

An Interdisciplinary Journal

ISSN: 1387-5868 (Print) 1542-7633 (Online) Journal homepage: www.tandfonline.com/journals/hscc20

Automatically generated place descriptions for accurate location identification: a hybrid approach with rule-based methods and LLM

Mikuláš Muroň, František Dařena, David Procházka & Roman Kern

To cite this article: Mikuláš Muroň, František Dařena, David Procházka & Roman Kern (2025) Automatically generated place descriptions for accurate location identification: a hybrid approach with rule-based methods and LLM, Spatial Cognition & Computation, 25:4, 269-311, DOI: 10.1080/13875868.2025.2449859

To link to this article: https://doi.org/10.1080/13875868.2025.2449859

Published online: 21 Jan 2025.

SPATIAL COGNITION & COMPUTATION
2025, VOL. 25, NO. 4, 269–311
https://doi.org/10.1080/13875868.2025.2449859

Taylor & Francis
Taylor & Francis Group

Automatically generated place descriptions for accurate location identification: a hybrid approach with rule-based methods and LLM

Mikuláš Muroň[a], František Dařena[a], David Procházka[a], and Roman Kern[b]

[a]Department of Informatics, Mendel University in Brno, Brno, Czech Republic; [b]Know-Center, Research Center for Data-Driven Business & Big Data, Graz, Austria

ABSTRACT
This paper explores the potential of using machine-generated descriptions to characterize a place in a way that humans can identify it. It presents a hybrid approach for generating place descriptions by combining rule-based generation of spatial relation facts with a LLM that converts these facts into natural language descriptions. The study focuses on urban areas and street-level scale, using OpenStreetMap as the primary data source. The rule-based method is informed by an experimental study that analyzed human-made place descriptions to understand reference object types used, their quantities, distances, and spatial relations. An experiment is carried out to assess the quality of machine-generated descriptions compared to human-made descriptions in a place identification task. The evaluation involved 70 participants identifying locations based on both human and machine-generated descriptions across a 200-hectare urban area. The results show that the same average identification accuracy was not achieved. However, the proposed method reached lower variance and the difference in accuracy is not substantial enough to impede place identification in the anticipated use cases. The method shows promise for applications in navigation systems, virtual assistants, and location-based services, particularly in situations where visual media cannot be used.

KEYWORDS
Spatial natural language; place description; place identification; OpenStreetMap; natural language generation

1. Introduction

Maps have been a fundamental aspect of human civilization for several millennia (Raaflaub & Talbert, 2009). However, maps in their current form are not the final frontier; they need to react and adapt to new trends and strive for even more seamless human-computer interaction (HCI). Natural language is considered one of the most intuitive forms of human communication. Although natural language interfaces are gaining popularity (Klopfenstein et al., 2017), they are not without challenges, bringing complexities that can

CONTACT Mikuláš Muroň mikulas.muron@gmail.com Department of Informatics, Mendel University in Brno, Zemědělská 1/1665, Brno 613 00, Czech Republic

lead to user frustration. Despite the existing challenges, it is a rapidly developing field with future potential. Recognizing this, Geographic Information Science (GIS) has identified the study of spatial language as an area of interest (Winter & Truelove, 2013), among which is the mediation of spatial information through language. Moreover, Weilenmann and Leuchovius (2004) argue that many location-based services would greatly benefit from the capability to mediate information in ways that are natural and relevant to users. The successful application of transforming spatial data into natural language is evident in various domains, including turn-by-turn navigation systems (Davis & Schmandt, 1989), description of geographic phenomena (Turner et al., 2009), generating of weather forecast (Reiter et al., 2005), photo captioning based on geographic context (Naaman et al., 2004, Hall et al., 2015), and others.

In the domain of natural language generation, large language models (LLMs) are increasingly being used instead of template-based or other traditional approaches (Bezirhan & Davier, 2023; Cook & Karakus, 2024; Michelet & Breitinger, 2024).

One specific area where such an approach appears particularly promising, but has not yet received enough attention, is the domain of place descriptions. People use natural language to describe places, and place descriptions are one of the most prevalent ways people convey information about space (Richter, Winter, et al., 2013). Expressions like "the hotel opposite the train station" are a natural way for people to refer to places when describing them (Herskovits, 1985).

Richter et al. (2012) define the concept of place in terms of how people perceive, memorize, reason, and communicate about space. For the purposes of our study, we adopt a much narrower definition of place. Specifically, we define a place as any identifiable point outdoors. As such, it typically does not have a specific name and must therefore be characterized with the help of its environment. The environment is a set of objects that can be identified on a map and are relevant to place identification. The scale of a place in our particular settings is fixed at street level, and only places in urban areas are considered.

1.1. Problem statement

Current approaches to mediate information about places primarily rely on visual representations, such as maps, to convey information (Balla et al., 2020). While maps and graphical interfaces are invaluable tools, they may not always be accessible or practical in every context. Considering how essential and ubiquitous spatial technology has become, it would be advantageous to be able to obtain information about a place even in situations where using visual media is impractical or unavailable, such as while driving or interacting with virtual assistants.

Despite the importance of natural language in conveying spatial information, the existing literature lacks a method that addresses the challenge of automatically generating a place description from map data. While studies have explored the transformation of spatial data into natural language in various domains, the topic of human-like and identifiable place descriptions remains unexplored. Most of the existing work focuses on generating short functional descriptions of places, such as photo captions, and does not address the challenge of finding the place using its description (De Carolis et al., 2009; Hall & Jones, 2022; Hall et al., 2015).

1.2. Research objective

This paper aims to bridge this gap by proposing a method that can automatically generate a natural language description of a place using map data while maintaining a similar level of identification accuracy as human-made place descriptions (also created from a map). The emphasis is on the human ability to accurately identify a place on a map based solely on its description. Furthermore, the proposed method should be able to produce naturally sounding descriptions of places in a way that is similar to how humans would do.

Our investigation is thus guided by the following question: *Can natural language descriptions of places located on a map – created by converting rule-based structured spatial representations into natural language using an LLM – achieve identification accuracy comparable to human-made place descriptions?* This question encapsulates the challenge of integrating natural language processing techniques with spatial relation facts. It is important to note that place descriptions created in this manner are based on map data rather than visual salience, an approach that has its own limitations. Additionally, while we design our approach to be as general as possible, we focus our efforts on one selected city.

1.3. Methodology overview

In order to automatically generate place descriptions, it is essential to gain an understanding of how people typically describe places. To achieve this, we conducted a study focusing on how people describe places in the given context and scope. We focus on understanding the quantity and type of features that people use when describing a place, with the aim of using these findings to inform a process of automatic description generation.

To accomplish our research objective, we define an empirically informed rule-based process that integrates all essential components: spatial data acquisition, relevance assessment, spatial relationship analysis, and generation of natural language descriptions. We leverage OpenStreetMap as a source of spatial data and a large language model for generating natural language descriptions of places.

We evaluate the validity of the proposed method by comparing the accuracy in the location identification task using descriptions generated by human participants and those generated by the proposed method. The results of this evaluation provide insight into the effectiveness of our method in accurately describing a location.

1.4. Potential applications

One potential application of the proposed automatically generated place descriptions is their integration into virtual assistants. Such integration appears to be a natural extension, given that virtual assistants are by their nature designed around a natural language interface. They are also frequently used on devices with no (or very limited) displays, such as smartwatches and smart speakers, yet often need to convey spatial information.

Another area where place descriptions could prove valuable is in turn-by-turn navigation systems, specifically, at the final stage, when the destination needs to be recognized. Merely announcing that the user has arrived at their destination, mentioning the relative orientation of the destination (for example, "Your destination is on the right"), or identifying it through simple landmarks might not always be effective in certain complex environments. It could be argued that a place description with a strong emphasis on identifiability can improve destination recognition (e.g., "Your destination is a low residential building with a crosswalk in front of it. Across the street is a subway station."). However, this approach may have limitations as we are utilizing mapping data to generate place descriptions that capture geographic attributes rather than visual salience.

Automatically generated place descriptions could also be utilized as part of assistive technologies to help build spatial awareness and conceptual models of unfamiliar environments. As discussed by Hakobyan et al. (2013), visual impairments pose challenges for space perception, while current assistive technologies have limitations in addressing these challenges. Place descriptions providing details about surroundings and spatial relationships between locations could aid space perception. However, further research is needed.

Another potential application lies in tourism, where generated descriptions of places can help users make decisions about what sites to visit. An illustrative example of such a system is presented by De Carolis et al. (2009).

2. Place description research landscape

There are many ways to convey information about a place in text form. A straightforward approach is to describe a place using coordinates or other forms of unambiguous identification (e.g., geohash or plus code). However, this is not the way people would usually use to describe or define a location. People are not able to easily comprehend a place described by computer-oriented identification (Weilenmann & Leuchovius, 2004). Even human-made place identifications (e.g., address, street name) require prior knowledge and don't provide much information about the characteristics of the place. According to Richter et al. (2012), the most natural way to describe a place is by using references to different kinds of urban landmarks, points of interest (POI), or even by involving personal perceptions and memories. Expressions like "the hotel opposite the train station" are a natural way for people to refer to places when describing them (Richter, Winter, et al., 2013).

According to Landau and Jackendoff (1993), a phrase describing a place requires at least three elements: figure object (object to be located), reference object, and their relationship. In general, we can distinguish three different ways of describing a place using a natural language: by place names (e.g., "Brno," "Graz"), deixis (context-dependent expressions like "here," "there"), or by contiguity/topology. Contiguity describes a place relative to nearby locations (e.g., "coffee shop next to the library"), while topology details its position within a larger arrangement (e.g., "house at the end of the street") (Levinson, 2003).

2.1. Characteristics of place descriptions

In their experiment, Zhou et al. (2005) collected and analyzed a large number of place descriptions to understand how individuals describe locations, identifying common types of place descriptions and key influencing factors of how these descriptions are tailored for different audiences. Also Richter et al. (2012) analyzed the predominant types of place descriptions. In a subsequent study, Richter, Winter, et al. (2013) investigated different levels of place granularity in place descriptions. Stock and Yousaf (2018) presented an interpretation of geospatial natural language utilizing a knowledge base of expressions. Their approach involves comparing novel expressions to those within the knowledge base to identify the most analogous expression and subsequently infer its meaning. Hall et al. (2011), Hall et al. (2015), and Hall and Jones (2022) have provided valuable insights through their detailed analysis of image captions. However, as this is a different domain from ours, applying all of their findings directly may present challenges.

2.2. Automatically generating place descriptions

The automated generation of descriptions has been receiving increasing attention. Dethlefs et al. (2011) present a method for generating adaptive verbal route descriptions in urban environments, incorporating salient geographic features and user familiarity. Their contextually sensitive approach has been validated through a user study. This work is directly relevant to generating place descriptions, as verbal routes thus generated (e.g., "Turn left into Grattan Street and follow it until you see the university on your right") exhibit parallels in structure and content to the place descriptions that our research aims to generate.

De Carolis et al. (2007) and follow-up De Carolis et al. (2009) introduce a system called MyMap, which is a mobile recommender system in the tourism domain. Among other things, this system generates comparative descriptions of places based on the user's interests and preferences. Their approach is built upon outlining the commonalities and differences among described places. Place properties are represented in a structured form (XML) and this representation is used to build the place description text from flexible templates.

Hall and Jones (2022) present a method that is in some aspects similar to our approach. They aim to create a system for generating complex, natural-sounding location descriptions. Their method is aimed at general documentation, which includes photo captioning, where events have taken place, etc. Thus, the generated descriptions are toponym-driven and descriptive (e.g., "Near the Volunteer's Walk in Old Town, City of Edinburgh"). This focus is also manifested in the dataset choice which is composed of labels from photo-sharing websites like *Flickr* (flickr.com) and *Geograph* (geograph.org). They introduce a novel toponym-driven algorithm for generating descriptions. However, we aim for a slightly different approach, emphasizing the identification and recognizability of a place based on its description on a map.

Another geospatial description generation system from which we can draw inspiration is presented by Hall et al. (2015). In this work, the authors introduce a system for generating captions for photos based on their geographic context. They analyzed existing photo captions and utilized the findings – such as the frequency of use of spatial prepositions and salience determination – to inform the design of a new system. This paper builds on the work of Hall et al. (2011), which focuses on the opposite process: interpreting spatial language in image captions. Although the domain of image captions is distinct from ours, certain analysis methods from these studies are relevant to our case and have served as valuable inspiration for our approach to place description analysis.

2.3. Data-to-text generation

It is plausible to expect that both the selection of appropriate features to generate place descriptions and the quality of the natural language presentation play roles in the effectiveness of these descriptions, i.e., a poorly constructed or unclear description would likely reduce the ability to accurately conceptualize the place being described. To mimic the style of human place descriptions, we take advantage of recent advances in natural language processing and streamline the problem by framing it as a data-to-text generation problem.

Automated text generation from structured data has attracted considerable attention in recent years (Lin et al., 2023), and a multitude of data-to-text generation techniques have been utilized in a wide range of applications (Axelsson & Skantze, 2023; Harkous et al., 2020; Kasner et al., 2023; Wang et al., 2018). The biggest advances in this area come from the recent development of large language models (LLMs) such as GPT-3 (Brown et al., 2020) or Bloom (Scao et al., 2022, preprint), which have caused a major paradigm shift. LLMs have been shown to outperform earlier methods in text generation tasks, showing a significant improvement in the quality and fluency of the generated texts (Li et al., 2022).

LLMs achieve notable performance by utilizing only natural language prompts and a few task examples as part of the input context (Gao et al., 2021). Although LLMs are not specifically designed for text generation from structured data, their versatility allows them to perform this task (Li et al., 2022). This versatility is demonstrated by the LLMs' performance in the few-shot learning mode (Wang et al., 2018). The text generation task is transformed into a language modeling task (Jin et al., 2023; Liu et al., 2023).

Zhao et al. (2023) provide a comparative analysis between LLMs and state-of-the-art models in the task of transforming structured data into natural language statements, showing that LLMs can produce text descriptions with higher fidelity to the original data. Additionally, Clark et al. (2021) demonstrate that human evaluations often cannot distinguish text generated by LLMs from human-authored text. While our method uses rule-based generation of structured spatial facts, we leverage LLMs to convert these facts into natural language descriptions. Although we do not explicitly evaluate stylistic quality, these findings suggest that LLMs can produce coherent and accessible descriptions.

3. How people describe places

In order to create human-like place descriptions that allow for place identifiability, it is essential to have an understanding of how individuals describe places in this context. While there is literature analyzing place descriptions, as discussed in Section 2.1, caution is necessary when extrapolating existing findings to our specific context – generating human-like place descriptions with identifiability as a primary consideration.

To inform the automatic description generation process presented in the following section, we need to address several specific subsidiary questions. To the best of our knowledge, no existing literature addresses these questions in this specific context and at the level of detail required. For this reason, we have decided to design an experiment that will serve as a precursor and whose results will be used to inform the process of automatic generation of place descriptions.

Q1) *Which type of reference objects (RO) do people tend to use when describing a place in an urban environment?* We aim to find out if people describe a location by mentioning a park nearby, a street name, a public transportation stop, a zebra crossing, or perhaps something else. We would also like to know if and how much people prefer certain types of RO over others.

Q2) *How many reference objects do people use when describing a place?* It would be beneficial to determine the number of RO in a place description that people consider sufficient for its effective identification.

Q3) *What is the typical range of distance between the described location and the reference objects?* With this question, we are trying to find out how large is the surrounding area that people take into account. Will they only consider the immediate vicinity or will they also refer to a distant reference object?

Q4) *What type of spatial relations do people use to describe a place?* Place description typically consists of a figure object, a spatial relation, and a reference object. We would like to find out which types of spatial relationships people prefer when describing a place (relative distance, direction, etc.) in a given context. We have adopted the classification scheme introduced by Landau and Jackendoff (1993) to categorize the types of spatial relationships.

However, we need to keep in mind that the answers to these questions are in some ways context-dependent. Contextual factors include the assumed recipient's prior knowledge, communication purpose and channel, roles of participants, and others (Garfinkel, 1967). If the context changes the description may also change. For example, when describing a place to be uniquely identified, a set of nearby objects and their relative locations will be included in the description. On the other hand, if we were describing the place to a friend, we might use subjective knowledge (e.g., "my favorite pub"). We can refer to the relevance principle claiming that the recipient will interpret the expression according to the relevant specific communication context (Richter, Vasardani, et al., 2013). The contextual setting of the experiment must therefore be consistent in certain aspects with the intended use.

3.1. Design of the experiment

We decided to conduct data collection in the form of a game. The use of location-based games has previously been proposed as a successful means of data collection. Winter et al. (2011) experimented with location-based mobile games as a tool for spatial knowledge acquisition called Tell-Us-Where. The advantage of this approach is the ability to set and shape the communication context. However, unlike Winter et al. (2011) who provided a very broad context (i.e., “Please tell us where you are”) or “Where are you?” used by Zhou et al. (2005), we would like to set the context more specifically. The purpose of providing a detailed context is to guide the user in such a way that the requirements and constraints we set are met (i.e., the place should be identifiable from the description, targeting only urban areas, no coordinate or reference system is imposed, and no subjective knowledge is used). For these reasons, we proposed the following location game assignment, which should sufficiently constrain the scope and define the context:

> Your task is to describe the place marked on the map to a colleague. Although they may not be familiar with the city, they should be able to find and recognize/identify this place based on your description. Don’t describe how to get to the place, but the place itself.

The location game is designed as a web application where the user is presented with a location on a map (depicted in Figure 1). The map is displayed

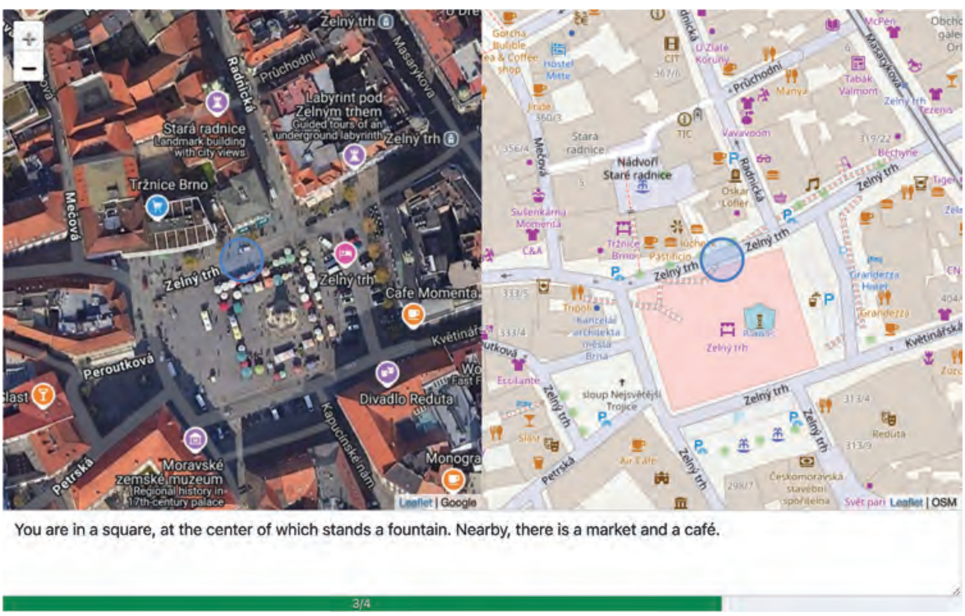

Figure 1. A screenshot from a web game for gathering place descriptions. On the left is the OSM basemap, on the right is the Google Satellite basemap. The blue circle indicates the location to be described. The text box shows an example of a real description written by a participant in the experiment.

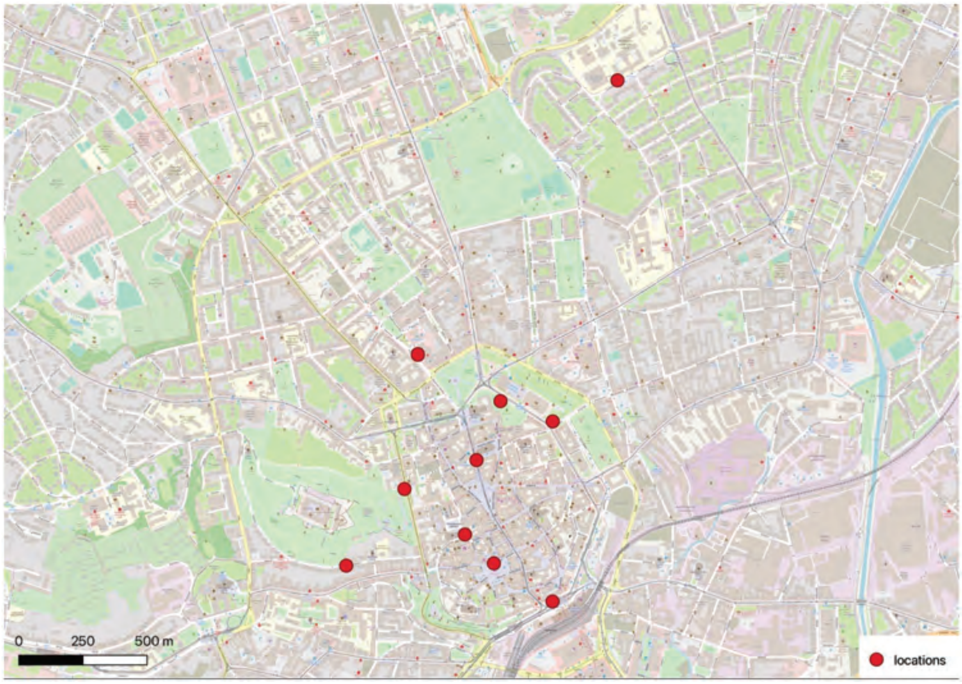

Figure 2. Depiction of the selected locations to be described. The locations were chosen to maximize diversity of urban typologies.

simultaneously using two different basemaps (satellite map and OSM basemap with default styling) to provide the user with a wider range of perspectives. The user is asked to enter a place description, while any typical interaction can be performed with the map (zooming, panning, etc.). In addition to the place descriptions, we also collected the following information: age, educational attainment, and acquaintance with the city. Users rated their acquaintance with the city on a four-point scale. We gather this information as it is plausible that the level of familiarity with a city might influence some aspects of a place description.

The decision to use maps, specifically OpenStreetMap (OSM) data, as a reference for creating place descriptions was made with the intention of using the same data source for subsequent automatic description generation as we aim to maintain a direct correspondence between this experiment and our approach to automatic generation of place descriptions.

Another important aspect of the experiment design was the choice of locations to be described. The scope of the experiment was geographically limited to an urban area of Brno, Czech Republic. A total of ten locations were selected (see Figure 2). Each user was tasked with describing four places, which were selected randomly from these ten locations. The number of places to be described by each user was chosen based on the trial experiment. It was observed that as the number of places increased, the attention and interest of the users decreased, resulting in a decrease in the quality of the descriptions.

When selecting the places to be described, an emphasis was placed on including a variety of urban typologies, as the characteristics of the area may influence the choice of reference objects. To represent the urban diversity of the area, ten places of different typologies were selected. Specifically, three places were selected in residential areas. The other three locations were situated in downtown areas, characterized by well-known landmarks. The remaining four places blended residential and commercial attributes, providing a diverse urban landscape. However, it is important to recognize that the character of the city itself introduces potential biases as urban morphology varies considerably between regions Oliveira (2021). This variation will influence the reference objects used and thus the descriptions of the places. The results and their applicability may be influenced by these differences, which should be taken into account when generalizing our results.

Participants were instructed to use their native language for the tasks, but the choice of language should not have a significant impact on our analysis, as the subsidiary questions were designed to be as language-independent as possible.

All procedures involving human participants were conducted in compliance with recognized ethical standards. All participants involved in this research provided their informed consent.

3.2. Analysis of human-made place descriptions

A total of 26 participants were involved in the experiment, each of whom was asked to provide four place descriptions, resulting in a total of 104 place descriptions collected, i.e., approximately ten descriptions per place. However, only 72% of place descriptions were considered valid. The rest did not match our quality criteria and were excluded from the dataset. The vast majority of participants were university students and academic staff. The 18–25 age group accounted for 92% of the collected data. This is also reflected in the educational attainment of the participants; the same percentage of people have completed secondary education. The analysis of the collected place descriptions was driven by our subsidiary questions introduced in Section 3.

We found that the descriptions collected from a given location were quite similar, as evidenced by the consistency in the types of reference objects (ROs) used. Specifically, the most common type of RO was present in an average of 74% of the descriptions for a given location, and the second and third most common types of RO were present in 54% and 41% of the descriptions respectively. This level of similarity in the descriptions may suggest that the number of descriptions collected effectively balanced the need for a sufficient sample size with the practical considerations of participant engagement.

Q1: Which type of reference objects (RO) do people tend to use when describing a place in an urban environment?

We manually extracted all ROs from the collected place descriptions. Any reference containing a toponym (e.g., "Mendel University") or an RO type (e.g., "bus stop") was considered RO. The ROs were subsequently assigned their alias according to the table by Dzial-Owski (2021) (e.g., Mendel University → university). In the compilation of collected place descriptions, a total of 215 references to RO were identified, encompassing 40 distinct aliases.

However, certain observed aliases exhibited thematic similarities. To enhance the generalizability of the results, we have grouped these aliases into relevant categories. Even if we did not observe an occurrence of some RO type

Table 1. Inferred thematic categories with corresponding OSM tags.

Category name	Observed RO type	OSM tags
Accommodation	hotel	building=hotel, tourism=hotel, hostel
Artworks and Monuments	statue, fountain	man_made=obelisk, historic=monument,man_made=statue, tourism=artwork
Culture	cinema,theater,art gallery	amenity=cinema, theatre, tourism=gallery, amenity=arts centre, tourism=museum
Education	school, university	amenity=college, university school, office=educational institution
Food and drinks	restaurant,fastfood, coffee, pub	amenity=bar, cafe, fast food, pub, restaurant
Greenery	tree, park, greenery	leisure=park,recreation ground, greenfield, grass, barrier=hedge, natural=tree, tree row, shrubbery, wood, tree=*
Office private	company hq, bank	amenity=bank office=company, financial
Office public	post office, government office	office=government, administrative amenity=post office building=government
Parking	parking house, parking spot, parking	parking=surface, amenity=parking, amenity=parking space, building=parking
Place of worship	church	amenity=place of worship, religion=*
Public Infrastructure	playground, stairs, gate, WC, zebra crossing	leisure=playground, highway=steps, crossing barrier=gate, lift gate amenity=toilets, water point
Public Spaces	district, street, square, street intersection	highway=*, place=square, district boundary=administrative
Public transportation	tramline,public transport stop, bus line, train station	public transport=*, railway=tram,subway,platform, tram stop, route=railway, train, tracks, tram, trolleybus
Shopping center	shopping center	shop=mall,department store, supermarket, wholesale
Shop	liquor store, bakery	shop=butcher, cheese, convenience, etc.
Sightseeing	castle, ossuary	tourism=attraction, building=castle, castle type=*

(e.g., greengrocery) in the collected place descriptions, we can assume that it belongs to a similar thematic category as RO types that occurred in the data (e.g., bakery), and thus has similar qualities.

It should be noted that what is referred to as an RO type (e.g., bakery) is typically represented as a tag within OSM data (e.g., shop=bakery). Consequently, each category is composed of features that correspond to a specific set of OSM tags. The categories derived from this process, along with the generalizations applied, are detailed in Table 1.

From the collected place descriptions, we aimed to identify which reference object (RO) types people prefer when describing a place. It is important to note that this approach does not directly assess the commonality or rarity of features in the environment. Instead, we focus on the choices individuals made when describing places, i.e., how likely participants would use a given RO type in their descriptions. By observing features not in the environment, but rather in the place descriptions, we analyzed reference objects that were actively selected for their descriptive utility, rather than their mere presence or frequency in the environment. This means that participants in our experiment effectively filtered out common and unimportant features by not including them in the descriptions.

We made a concerted effort to select a diverse range of locations, including their respective surroundings. However, since it was not possible to predict which types of ROs would appear in the descriptions, occurrences of some RO types may be misrepresented. Therefore, we propose a metric that considers both local occurrences (in the descriptions of one place) and global occurrences (in the descriptions of all places). We term this metric *Observed Descriptive Utility* (ODU), which ranges in value from zero to one and is calculated according to the following methodology: The reference object type is deemed to be present in a location if its relative frequency reaches at least 5%

Table 2. RO types with their corresponding observed descriptive utility value.

RO type	Observed Descriptive Utility
Public Spaces	100%
Greenery	60%
Food and Drinks	60%
Public Transportation	50%
Public Infrastructure	40%
Culture	30%
Shop	30%
Parking	20%
Place of Worship	20%
Artworks and Monuments	20%
Shopping Centre	20%
Office private	20%
Office public	10%
Education	10%
Sightseeing	10%
Accommodation	10%

considering the types of the reference objects appearing in all descriptions of that location (a description may contain several reference objects, even of the same type). The ODU for a given RO type is then calculated as the ratio of the number of locations where the RO type is present to the total number of locations. This prioritizes RO types that occur in multiple locations over those that occur frequently in only a few locations. For example, the mention of "greenery" in three different place descriptions from three different locations is considered more significant than the mention of "greenery" in nine descriptions from one location.

Table 2 presents the results of this process. It contains the value of ODU, i.e., the percentage of locations containing that specific RO type for each RO type, classified according to Table 1.

We acknowledge that the presented metric is influenced to some degree by the availability of features in the given area and thus does not capture geographic features across all possible places. Therefore, it would be necessary to perform the ODU computation again for a city with different features and availability to better match and assess the descriptive utility of those features in varied contexts.

As might be expected, references to public spaces, including districts, streets, and squares, were consistently employed across all the locations studied. Utilizing these elements to describe a location is intuitive for individuals, especially when they have a basic familiarity with the area (Kuipers, 1978). Green spaces emerged as the second most common category in place descriptions, along with the Food and Drinks category.

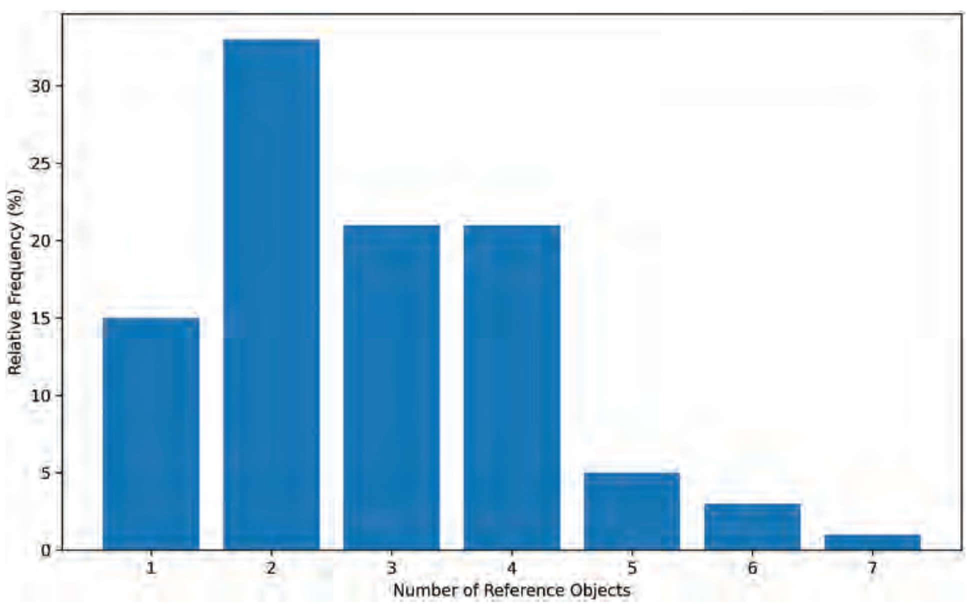

Figure 3. Distributions of reference objects frequencies in place descriptions.

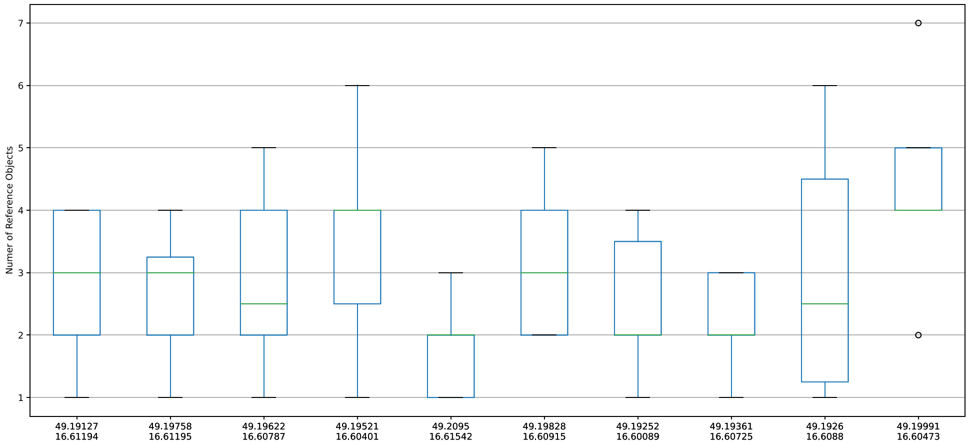

Figure 4. Distribution of occurrences of referenced objects in place descriptions for each location.

Q2: How many reference objects do people use when describing a place?

For each collected place description, we quantified the number of reference objects contained. For instance, the description "The is a place near a *hotel* by the main *train station*" corresponds to two ROs. All descriptions contained at least one RO, which may indicate a well-set context (i.e., the participants in the study demonstrated an understanding of the task at hand).

The distributions of the numbers of RO observed in descriptions are shown in Figures 3 and 4. We have generally observed two types of strategies employed by users when describing a place. The first and more common is highly descriptive, where the user begins by outlining a broad surrounding and then refines the location with additional, more specific ROs. This approach typically results in a higher number of ROs being used. Conversely, the second strategy involves providing one or two highly specific locative expressions containing very significant ROs without any additional context (e.g., "in front of a well-known building"), resulting in descriptions that are concise and less detailed.

Q3: What is the typical range of distance between the described location and the reference objects?

The distance to the location was measured for each RO mentioned in the description. In cases where the OSM feature corresponding to an RO was a polygon, the shortest distance to it was measured. In the case of RO containment (i.e., the location is inside an RO) the value zero was set as the corresponding distance. The data showed that the distance between the figure

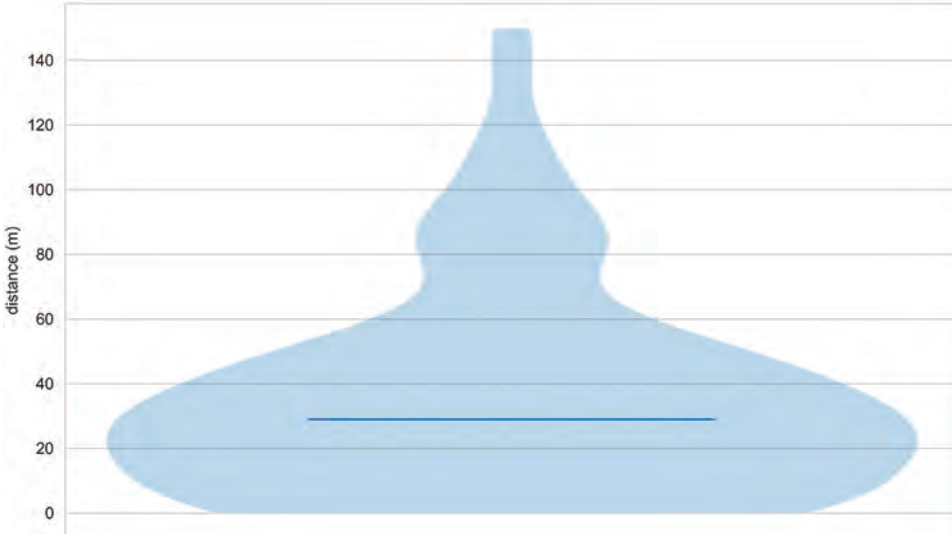

Figure 5. The distribution of relative distances from figure object to RO, displaying data up to the 90th percentile.

object (object to be located) and the RO ranges from 0 to 1100 m with the median at 30 m (see Figure 5 for more details).

We conclude that, in the given context, people predominantly utilize ROs situated within a 154 m radius (representing the 90th percentile) to describe a place. For subsequent analyses, the data was truncated at this 90th percentile.

It's important to note that the observed distances stemmed from the use of the OSM map as a reference when creating place descriptions. This choice was driven by the need for data source consistency between human-made place descriptions and their automatically generated counterparts. It is plausible that place descriptions created from different references (e.g., in-situ, panoramic photo) could yield different results.

We recorded how users interacted with the map when describing a location, i.e., whether the default map view presented was appropriate in the given context. The application placed the object to be described in the center, approximately 100 meters from the map boundary. We recorded zoom, pan, and drag actions. Based on these measurements, we found that users performed at least one action on the map in 52% of cases, while the most typical action was zooming. These findings indicate that more than half of the users

Table 3. The relative frequencies of spatial relation types observed in the dataset.

Spatial relation type	Relative frequency
relative distance	43%
direction	32%
paths or trajectories	25%

found it necessary or beneficial to adjust the default map view to better describe a location.

Q4: What type of spatial relations do people use to describe a place?

We identified 188 spatial relationships in our dataset, corresponding to an average occurrence of 2.5 in each place description. These relationships were manually classified into broad categories, as introduced by Landau and Jackendoff (1993). The relative frequencies of the various types of spatial relations are presented in Table 3.

The most frequent were spatial relationships related to relative distance (near, on, far, etc.). Landau and Jackendoff (1993) define this category as the expression of the distance between a figure object and a reference object. It is interesting to note that we did not observe a single occurrence of an expression of the absolute distance in the dataset (e.g., “100 m from here”).

The second most common category of spatial relationship is direction. We have observed a total of seven unique direction relations. The most common relations in this category were: opposite (18/188), next to (13/188), and in front of (12/188); the other ones had a frequency of five or less. The least numerous type were linguistic expressions defining paths or trajectories. This category specifies motion or orientation (via, toward, away from, etc.). People strongly preferred axis-oriented types (forward, ahead, etc.) and operators on regions (into, toward) over earth-oriented types (east, west, etc.) (40/188 vs. 2/188). It is worth noting how rarely people used cardinal directions.

4. Generating place descriptions

The main objective of this section is to address the question of whether natural language descriptions of places located on a map can be automatically generated in a way that matches human-made descriptions in terms of identification accuracy. To tackle this challenge, we propose a method that integrates an empirically informed rule-based process with natural language processing techniques. Our approach is informed by an understanding of how humans describe places, which we established in the previous section. Leveraging these insights we aim to develop a generation technique that can automatically generate natural-sounding place descriptions that have comparable identification accuracy to those produced by humans.

Given the diversity of place descriptions, narrowing the scope is necessary to draw conclusions. Accordingly, the scope and boundaries of the proposed method are delineated as follows: we focus on urban areas, specifically those with characteristics similar to the urban area where the case study was conducted. This focus is motivated by the diverse range of geographic features found in urban environments, which offer a rich source of information for

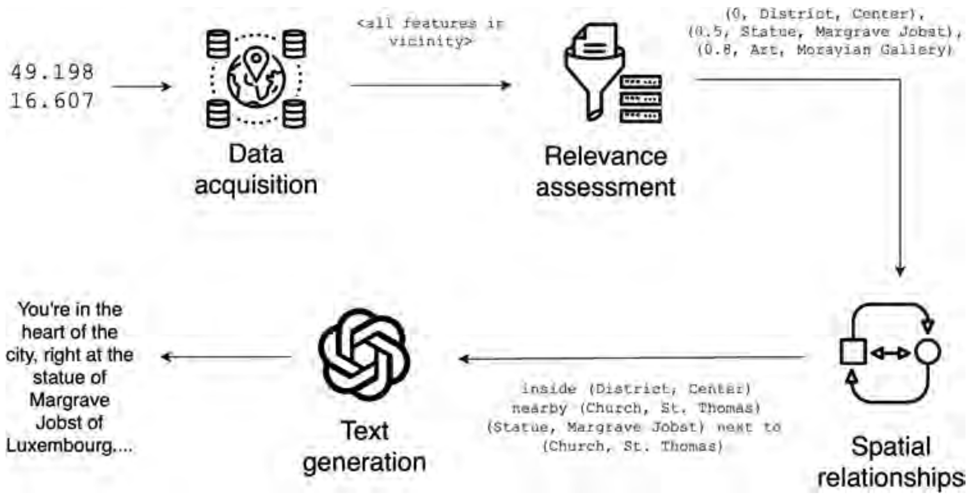

Figure 6. Steps of the proposed method for automatic generation of place descriptions.

generating informative place descriptions. Additionally, the demand for detailed descriptions of urban places is likely to be higher given the high concentration of individuals and businesses in these areas. We also require that the place in question is identifiable by its description by an uninvolved actor, i.e., any subjective knowledge about a place is not taken into account (e.g., "I'm at the bar where I broke my leg"). This requirement also includes any subjective emotionally tinged descriptions (e.g., "I'm at my favorite place.") or idiosyncratic descriptions (e.g., "It's the pub we went to last time"). Moreover, only relative and intrinsic frames of reference are considered (i.e., no coordinate or reference system is imposed). To facilitate coherent conclusions, our automatically generated place descriptions adhere to the same geographic scale, specifically the street level, as employed in our prior experiment described in Section 3.

4.1. Method overview

The proposed method consists of four steps, as illustrated in Figure 6. The first step involves obtaining spatial data in the vicinity of the figure object (a place to be described). The next step is ranking, which aims to evaluate the features obtained using a score that expresses their relevance. Only n features with the highest scores are selected for subsequent processing. The value of n is set to 5, as discussed in Section 4.6.2. In the third step, the spatial relationships between features and the figure are identified and explored. The objective of these steps is to acquire and extract the relevant data necessary for generating natural language. This acquired data is then transformed into intermediate representations. By intermediate representations, we refer to a data format, in the form of tuples, which facilitates the integration of structured knowledge into the model. Such a representation appears to be more suited to LLMs than more sophisticated structured formats and is therefore expected to produce better results in text generation tasks. We draw our inspiration from Moiseev et al. (2022) that used a similar approach. The last step is to transform an intermediate representation into a natural-sounding and coherent description of the place. In this phase, we reframe data-to-text generation tasks as a language modeling task and employ a large language model, specifically GPT-3 (text-davinci -002), to produce place descriptions. We aim to harness the model's text-generation capabilities, sidestepping its potential for internal reasoning and factual knowledge. Such a strategy aims to produce place descriptions grounded in the input data, avoiding the pitfalls of hallucination and enhancing output consistency. We do not claim that using an LLM directly improves identification accuracy, it is just one step in the process of creating natural-sounding place descriptions.

4.2. Spatial data acquisition

It is not feasible to create a place description without spatial data of sufficient quality. The OpenStreetMap platform has been widely utilized in comparable scenarios and contains all the necessary data for our use case.

OpenStreetMap, being crowdsourced data, inherently carries a risk of varying data quality. In this context, quality refers to the accuracy, completeness, and timeliness of the data available. To assess quality, we utilize intrinsic quality metrics like *Editing Density* (the number of edits within a certain area over time), suggested by Minghini and Frassinelli (2019). In our case study, Brno shows 18k edits (Anderson, 2016), comparable to similar-sized European cities, indicating satisfactory data coverage and, therefore, presumed quality.

The first step in generating a place description is acquiring relevant spatial data in the form of features. Within the platform of OpenStreetMap, any physical element that can be mapped is considered a feature (Arsanjani et al., 2015). To acquire OSM features, we decided to use the widespread *Overpass API*. Queries are formulated using the *Overpass Query Language*.

A total of two queries were made. The first one was hierarchical and its purpose was to retrieve the hierarchical arrangement of the containment features, i.e., to put the given location into a spatial hierarchy (e.g., "A place is in the park which is in a neighborhood which is in a district").

The second query retrieved all feature types (nodes, ways, relations) from the vicinity of the given place. The maximum distance was set to 154 m, corresponding to the 90th percentile of an RO distance range observed in our dataset. Any features located within the determined distance were incorporated into the result, regardless of whether they were fully contained within the area. The maximum distance parameter can be adjusted, allowing our method to be easily used in other levels of hierarchy (e.g., to describe a district).

4.2.1. Data preprocessing and normalization

Each OSM feature (=physical feature present on the ground) is described by multiple tags (e.g., amenity:townhall, building:public). An overview of the primary feature types used is available at OSM Contributors (2023). However, the key and value combinations in tags are standardized only informally and their categorization into primary feature types (e.g., Amenity, Barrier, Building, Craft, Emergency, Geological) is not suitable for our purposes. For this reason, all tags were transformed into human-centric aliases using a translate table provided by OSM Scout Server (Dzial-Owski, 2021) (e.g., amenity=cafe → Cafe, amenity=waste_basket → Bin). This transformation simplified interactions with the data and enabled their integration into the thematic classification scheme established in Section 3.2.

Features were selected for further analysis only if they were tagged with labels that were a part of the thematic classification scheme. For instance, even though the dataset did not have an entry for a butcher (tagged as shop=butcher), it will still be included in further analysis because it falls under the broader *shops* category that is part of the classification scheme. Conversely, a paleontological site (tagged as geological=palaeontological_site) will not be included in further analysis because it was not only absent in the dataset but also did not align with any thematic category outlined in Section 3.2. The rationale for the exclusion of such features is that it is not possible to assess their relevance if they have not been observed. We acknowledge that such an approach may occasionally omit features that could have some relevance. However, given the abundance of features available at any given place, we prefer to prioritize features for which there is a prior observation rather than not, as we do not want to dilute the place description with potentially irrelevant features.

Another tag that plays a significant role is name, along with its variations loc_name, official_name, etc., as these tags hold the place name. However, it is important to note that not all features are associated with names.

From the obtained data, we keep for each OSM feature only its coordinates, alias, and name (if one exists). The resulting set of elements then serves as the input for the subsequent relevance assessment.

4.3. Relevance assessment

In the previous section, we have described the method for obtaining features from the vicinity of a figure object. The number of features obtained in this manner is in the order of higher tens up to lower hundreds (depending on the area). Incorporating all features as reference objects is not feasible, as this approach would result in a description that is unreasonably lengthy and predominantly filled with non-relevant features. Hence, we need to filter ROs and keep only the relevant ones. When assessing the relevance of a feature, we consider two criteria: distance (the closer the feature is to the figure object, the more relevant it is) and the Observed Descriptive Utility metric presented in Section 3.2.

4.3.1. Distance

From the preceding step, we have a set of features represented by a triple (coordinates, alias, name). We used the Euclidean distance between the figure object and the feature as our distance measure. Although there are other metrics for expressing distance, such as walking distance, their utility, in this case, is deemed minimal and were, therefore, not utilized.

In the last step, the calculated distance was normalized to a zero-to-one range to represent a relative distance. Value one represented the maximum possible distance, which was 154 m in our case (90th percentile of the maximal observed distance). However, this value can be readily adapted for alternative configurations. The output of this step was a set of features characterized by the following triples: (relative distance, alias, name), for example, (0.5, statue, Margrave Jobst of Luxembourg).

4.3.2. Relevance score

The objective of this step is to evaluate features based on their relevance, with the intent of filtering out irrelevant ones. The relevance score is determined using the Observed Descriptive Utility (ODU) metric (refer to Table 2) and the relative distance of a feature from the figure. The relevance score is computed using the following formula:

$$\text{relevance score} = (1 - d^2) \cdot \text{ODU}$$

where d represents the normalized relative distance. People are generally more interested in what is happening near them than what is happening far away (Walmsley, 1974), this aspect is captured in the term d. Moreover, people are more attentive to specific types of features over others which is captured by *ODU*.

The relevance score is calculated for all available features, the results are sorted, and exactly the top n candidates are selected for a description generation. We set the value of n to five. As we have observed in Question 2 Section 3.2, people use two different strategies when they describe a place (referencing a few prominent ROs or a very descriptive explanation referencing many

ROs). We adopted the second strategy as we expect this approach to be more accurate and robust. This decision is further supported by the findings of Hall and Jones (2022), who confirmed their hypothesis that people prefer more detailed location descriptions.

Although the number of ROs in human-made descriptions varied from place to place we use a fixed number of ROs. This strategy aims to reduce potential identification issues by always utilizing more ROs. This minimizes the risk of selecting an incorrect (i.e., unimportant) RO, thereby increasing the likelihood of successful place recognition.

4.4. Spatial relationships

Spatial relationships form the cornerstone of our understanding and conceptualization of any physical environment. Thus, it can be argued that place descriptions that include spatial relationships have the potential to be more precise, providing specific details about where objects are located relative to each other.

We categorize spatial relations into two broad categories: relations representing relative distance (how far the figure object is from the RO) and directional spatial relations (how the RO is positioned relative to other ROs or the figure object).

4.4.1. Relative distance

First, we focus on the relative distance as it emerged as the most frequently observed spatial relationship category. Whilst maintaining the same relative distance relationships classification as Landau and Jackendoff (1993), i.e., in the interior, in contact, proximal, distant, we have decided to subdivide the proximate category into two subcategories, aiming for a more nuanced separation. This decision leads to the introduction of *Immediate Proximity* and *Close Proximity*.

- *Category 1)* In the Interior (e.g., in, inside)
- *Category 2)* Outside, but in Contact (e.g., on, against, next to)

Table 4. Distance categories with corresponding relative and absolute distance ranges.

Category	Relative distance	Absolute distance (m)
1	0	0
2	0–0.03	0–5
3	0.03–0.2	5–30
4	0.2–0.7	30–105
5	0.7–1.0	>105

- *Category 3)* Immediate Proximity (e.g., right at, very close to)
- *Category 4)* Close Proximity (e.g., near, nearby, close to)
- *Category 5)* Distant (e.g., far, in the distance)

Category 1 (e.g., in, inside) is used in cases when the described place (figure) is located inside the RO (e.g., in a district, inside a park). In such cases, we utilize data from the hierarchical OpenStreetMap Overpass query, as described in Section 4.4. *Category 2* (e.g., on, against, next to) includes cases where the buffer (5 m) surrounding the figure intersects with the RO. *Category 3* captures the spatial relationships where objects are close to each other, almost touching, while *Category 4* is used to describe scenarios where objects are near each other but with a discernible separation. *Category 5* applies when the described place is significantly separated from the reference object, often beyond immediate reach.

Relative distance categories are then mapped to normalized relative distances, as detailed in Table 4. The thresholds for these distance categories were derived from the observed distribution of distances in the collected place descriptions from the experiment described in Section 3. For *Category 1*, a value of zero indicates containment within an area. The boundary for *Category 2* is set at the largest of the frequent small nonzero values observed, which is 5 meters (0.03 normalized). The range for *Category 3* extends from this point up to the median. *Category 4* spans from the median distance to the 85th percentile, while *Category 5* encompasses the distances beyond this point.

4.4.2. Direction

In our analysis, we have selected the three most prevalent directional spatial relations observed, as outlined in Question 4, Section 3. These include *opposite* (accounting for 38% of instances), *in front of* (26%), and *behind* (10%). The following rules apply to the definition and application of these relationships:

- **opposite**: RO_1 is opposite to RO_2 and vice versa if the ROs are approximately equidistant from the figure object, there is no other RO in front of them and a line can be constructed between RO_1, the figure, and RO_2 with a maximum divergence of 15°.
- **behind/in front of**: RO_1 is behind RO_2 when the relative distance of RO_1 is significantly greater than that of RO_2 and when the angle between the

directions to RO_1 and RO_2 from the figure is less than 30°. The opposite of this spatial relation is considered as in front of.

When defining the spatial relationships above, we were constrained by the need to define them precisely to facilitate their algorithmic detection. These operationalizations were inspired by their use in observed human-made place descriptions. While we have not conducted exhaustive testing of these parameters, we emphasize the importance of future research to refine these operationalizations.

4.5. Intermediate representations

This section aims to present an intermediate representation, which encapsulates all the vital information needed to characterize a place in a structured format. Data with a clear and simple structure can be easily understood and processed by LLMs, enabling the models to learn and generate accurate outputs (Moiseev et al., 2022). Moreover, the records in the presented representation are interconnected, forming a network of related facts. This can help LLMs understand the context and relationships between different ROs, which should in principle lead to more coherent and contextually accurate place descriptions.

To create the intermediate representation, selected relevant features represented as tuples consisting of the alias and optionally the feature name were supplemented by the distance or directional relationships. The intermediate representation structure represented in the Backus-Naur Form (BNF) is as follows:

```
<result> ::= <relation_distance> ” ” <RO> |
             <RO> ” ”
<relation_directional> ’ ” <RO>
<RO> ::=
<relation_distance> ::= (”in‘ | ’inside”) |
                        (”on‘ | ’against”) |
                        (”right at‘ | ’very close to ”) |
                        (”near‘ | ’nearby‘ | ’close to”) |
                        (”far‘ | ’in the distance”)

<relation_directional> ::= ”next to” |
                           (”behind‘ | ’in front of”) |
                           ”opposite”

<ro_name> ::= [a-z]*

<ro_type> ::= [a-z]*
```

An example of an intermediate representation of a place with five ROs used in five distance and two directional relationships is presented below:

```
inside (District, Center)

right at (Statue, Margrave Jobst of Luxembourg)

nearby (Church, St. Thomas)

close to (Public transport, Česká)

in the distance (Art, Moravian Gallery – Governor’s Palace)

(Statue, Margrave Jobst of Luxembourg) next to (Church, St. Thomas)

(Art, Moravian Gallery – Governor’s Palace)behind (Statue,
Margrave Jobst of Luxembourg)
```

4.6. Data to text generation

This section introduces a way to generate a natural-sounding description of a place based on its intermediate representations. We transform and reduce the problem of generating a place description to a problem of data-to-text generation. As outlined in Section 2.2, we chose the LLM-based approach to accomplish the task given its superior performance as demonstrated in the existing literature (Chung et al., 2023). Moreover, it has been shown that texts generated by LLMs are indistinguishable from human-authored text (Clark et al., 2021; Karpinska et al., 2021; Radford et al., 2019). Thus, such an approach could produce place descriptions that are both stylistically and linguistically comparable to those produced by humans.

Building upon the findings of Keymanesh et al. (2022), Nan et al. (2021), and Mosbach et al. (2023), we opted to employ a blended method, which involved fine-tuning the model on a limited set of data and setting the right context through a few-shot prompt initiation. In this study, we opted to use the GPT-3 (text-davinci-002) model introduced by Brown et al. (2020), which was recognized as one of the leading models available at the time of our experiment. This decision was based on its prominence in several areas of natural language processing.

4.6.1. Prompt design

An important aspect when using language models for text generation is the design of a prompt, as PLMs are sensitive to prompt design and improperly crafted prompts lead to poor performance (Arora et al., 2022; Gao et al., 2021). The area of prompt engineering is extensively researched. In line with the recommendations of Liu et al. (2023), our prompt design employs a prefix label structure. The prefix label data introduces an intermediate structured representation. Conversely, the prefix description conveys the corresponding place description in natural language. An illustrative prompt is presented below:

```
data:

inside (District, Center)

right at (Statue, Margrave Jobst of Luxembourg)

nearby (Church, St. Thomas)

close to (Public transport, Česká)

in the distance (Art, Moravian Gallery – Governor’s Palace)

(Statue, Margrave Jobst of Luxembourg) next to (Church, St. Thomas)

(Art, Moravian Gallery – Governor’s Palace) behind (Statue,
Margrave Jobst of Luxembourg)


description:
You’re at the heart of the city, right by the statue of Margrave
Jobst of Luxembourg. Adjacent to the statue stands St. Thomas’
church, with the Moravian Gallery located just behind it. Close
to this area, you’ll find the Česká public transport station stop.
```

The prompt consists only of data and description sections with no guiding information. Our experimental observations showed that variations in prompt strategies, such as adding different variants of explicit task instructions (i.e., *Your task is to generate descriptions of places so that the place can be recognized*, etc.), mainly affected the stylistic aspects of the generated text without significantly affecting its overall quality or informational content, and with variation in most cases limited to different wording. While further comparative analysis of prompt variations could provide additional insights, such an exploration is beyond the scope of the current study.

To establish the correspondence between the intermediate representation and the place description, a dataset was constructed. This dataset was intended to illustrate the process by which humans create a natural language place description from an intermediate representation. A sample of four participants was recruited with the task of composing place descriptions with the same goal as in the previous experiment but were only given the intermediate representation. This resulted in a collection of 19 place descriptions – intermediate representation pairs. After quality assessment, one pair was removed due to poor quality. Three of the remaining 18 pairs, following the few-shot approach Min et al. (2022), were used as part of the prompt, while the rest were utilized to fine-tune the model.

location: 16.6152, 49.1831
typology: brownfield
description: You are located in the Trnitá district on Rosická Street at the bus stop Dolní nádraží. On the same street, you are close to a railway track and a bit further there is a park and a parking. In a greater distance, there is the railway station Brno dolní nádraží.

location: 16.6034, 49.1973
typology: city center
description: This is a place in the city center. Nearest to you is the JA Komenského church. Across the crossing with traffic lights, there is the courthouse Ústavní soud České republiky. A bit further there is Fakulta sociálních studií Masarykovy univerzity. In a greater distance are the Joštova and Opletalova streets with the public transport stop Komenského náměstí.

location: 16.6007, 49.1918
typology: semi-residential area
description: You are near the Pekařská street and a greenery. Next to you there is a restaurant Shanghai. Further away there is a hospital with the name Fakultní nemocnice u sv. Anny and also the park Špilberk. Nearby there is the public transport stop Nemocnice u sv. Anny.

Figure 7. Examples of automatically generated descriptions of three places with different characteristics.

4.6.2. *Fine-tuning*

Fine-tuning enhances the few-shot learning approach by exposing the model to a larger number of examples beyond the prompt capacity, resulting in improved performance across a diverse set of tasks (OpenAI, 2023a). We used 15 pairs of intermediate representations with corresponding place descriptions for fine-tuning via the provided API. The learning parameters (n_epochs, batch_size, etc.) were left at the values recommended by OpenAI (2023a). While the sample size is quite limited, it should be noted that the fine-tuning process serves as supplementary and is not crucial for the model's performance. Our experiments indicate that the model achieved comparable results solely based on the provided prompts even without fine-tuning (i.e., few-shot).

4.6.3. Inference

During the inference phase, the model is provided with the prompt (refer to Section 4.6.1), which contains the intermediate representation of the place to be described. The prompt is terminated by the prefix label description: (i.e., an instruction to generate a description). The model continues to generate text until either a maximum of 2,048 tokens is produced or a stop sequence, represented by a new line, is encountered. The latter behavior is matched to the structure of our training data, as in our dataset each description block ends with a new line, instructing the model to stop generating text when a description segment is complete. This approach ensures consistency with the learned format. Regarding the 2,048 token limit, this threshold was set as a hard limit, as descriptions beyond this length would be excessively verbose.

During the experiments, we utilized a *temperature* hyperparameter of 0.2 while maintaining the *top_p* value at its default setting of 0.9. This decision aligns with the recommendations of OpenAI (2023b), who advise modifying either the *temperature* or *top_p*. The choice of a temperature value was informed by our empirical observations regarding its influence on the model's output. A higher temperature value yielded more diverse output but increased the model's tendency toward excessive creativity, potentially leading to deviations from the provided intermediate representation data. To counteract this, we opted for a lower temperature setting, ensuring that the generated output remains closely aligned with the input data. This is consistent with the observations of Bhavya et al. (2022) who achieved the best results with precise imperative prompts and low-temperature settings.

Figure 7 presents illustrative place descriptions of previously unseen locations generated by our method. These locations represent samples from various urban typologies to demonstrate the versatility of our method across different contexts. The presented descriptions are not selectively curated.

5. Evaluation

This section describes the evaluation process of our proposed method for generating place descriptions, focusing on the assessment of identification accuracy, i.e. whether and to what extent it is possible to find a place on a map using only its description. This analysis is needed in order to understand the practical applicability of the proposed approach in real-world scenarios. We have considered several benchmarks to compare identification accuracy.

While LLMs embed some geographical knowledge and can generate place descriptions solely on the internal knowledge derived from training data, the accuracy diminishes significantly at more granular levels. Based on our experiments, the place descriptions generated using only model internal knowledge contain too many factual inaccuracies to be used for comparison.

To overcome this issue, we have turned our attention to an LLM, which uses an external data source for knowledge. Google Bard, powered by PaLM2, can access spatial data by utilizing the Google Maps extension. However, despite trying various prompt variants (including instructions given to the participants of our experiment), the results mixed actual facts with hallucinations, making them unusable to be considered for comparison. In the appendix, we provide some of the best results we were able to obtain for comparison with our method's results (which, unlike Google Bard's, are not cherry-picked).

Another possibility is to use descriptions from systems originally developed for purposes other than place identification (such as Hall and Jones, 2022). However, as these systems are not optimized for this specific task, their use as a benchmark may be inappropriate for fair comparisons. To the best of our knowledge, there is currently no research that we can directly compare our results to, as no other studies have focused specifically on the generation of place descriptions from map data with identifiability as the primary criterion. With these considerations in mind, we have chosen to benchmark against human-made place descriptions.

5.1. Evaluation method

As we aim to generate descriptions that facilitate place recognition and identification, the core principle of the evaluation is to measure and compare the accuracy of a place identification on a map between the place description generated by our method and the reference place descriptions made by humans.

Seven locations, chosen from the same collection utilized in the previous experiment, were selected for comparison in this manner. For each location, we had a description generated by our method and also three selected descriptions written by humans from the previous experiment described in Section 3. All place descriptions used in this experiment for both the human descriptions and the descriptions generated by our method are presented in the Appendix.

The evaluation was conducted as an experiment using a web application. The subject's task was to mark the location on the map according to the given place description, not knowing whether the description came from a human or a computer. The application then recorded the difference (as a geometric distance) between the actual location and the guessed location.

When selecting locations, we aimed for diversity, choosing two locations in residential areas, three in the downtown, and two that mixed residential and

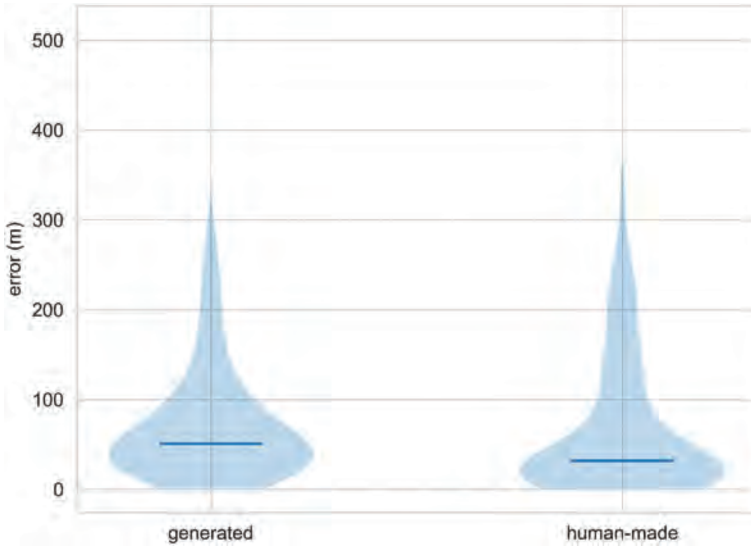

Figure 8. Errors in place identification using a place description made by humans vs. generated by our method.

commercial features, offering a varied urban landscape. However, we acknowledge that the city's character may introduce biases, as urban morphology can differ across regions (Oliveira, 2021).

The sample size of seven was chosen to balance the desire for a sufficient number of samples with the need to reduce the burden on participants. Whether the user was shown one of the human-made descriptions or a machine-made one was determined by chance.

The area for locating a place based on the description (i.e., the area to which the map was first zoomed) was set as a bounding box (with a small buffer) containing all seven places included in the study. The size of the search area thus created was approximately 200 ha.

5.2. Study settings

The participants of the experiment were similar in composition to those in the previous experiment. The data collected reveals that 90% of the participants belong to the 18–25 age group. This demographic is also reflected in the educational attainment of the participants, the same percentage of people have completed secondary education. We put in place measures to eliminate the possible overlap of participants between this experiment and the earlier experiment detailed in Section 3. The number of participants was larger than in our previous study. In total, we collected 490 samples from 70 participants.

Each sample consisted of a guessed location from which the error, i.e., the distance between the location guessed from the provided place description and the actual location, was calculated. In the experiment, each participant was asked to guess the location of 7 places, resulting in a total of 70 samples for each place. The users' guesses were based on two types of descriptions: approximately half of the samples were based on human-made place descriptions, and the other half were based on place descriptions automatically generated by our method.

The vast majority (80%) of participants were at least somehow familiar with the area over which they identified locations (a questionnaire was conducted to assess the subjects' familiarity with the study area).

5.3. Evaluation results

The violin plot in Figure 8 depicts the distribution of identification errors across all locations. The median identification error was 34 meters for human-made descriptions and 54 meters for machine-made. The mean identification error was 75 meters for human-made descriptions and 70 meters for machine-made descriptions.

More indicative than the absolute value of the identification error is the accuracy compared to human descriptions. Although automatically generated descriptions were identified with a 20-meter higher median error compared to human-made descriptions, this difference is not substantial enough to prevent the identification of a location at the street-level scale we are targeting for most of the anticipated use cases described in Section 1.1.

Interestingly, the identification errors for automatically generated descriptions had a lower variance than those for human descriptions. A plausible explanation for this observation is that automatically generated descriptions exhibit greater consistency, stemming from a singular process, whereas the quality of human descriptions varies with the competence of the author.

In conclusion, automatically generated descriptions exhibited a greater systemic error, as indicated by a higher median, yet demonstrated increased consistency, as represented by lower variability. These two phenomena subsequently lead to nearly identical arithmetic means observed for both categories.

Variations in the results were expected across different locations. For two of the seven locations, the median error in place identification favored the descriptions generated by our method. Detailed breakdowns for each location are illustrated in Figures 9 and 10.

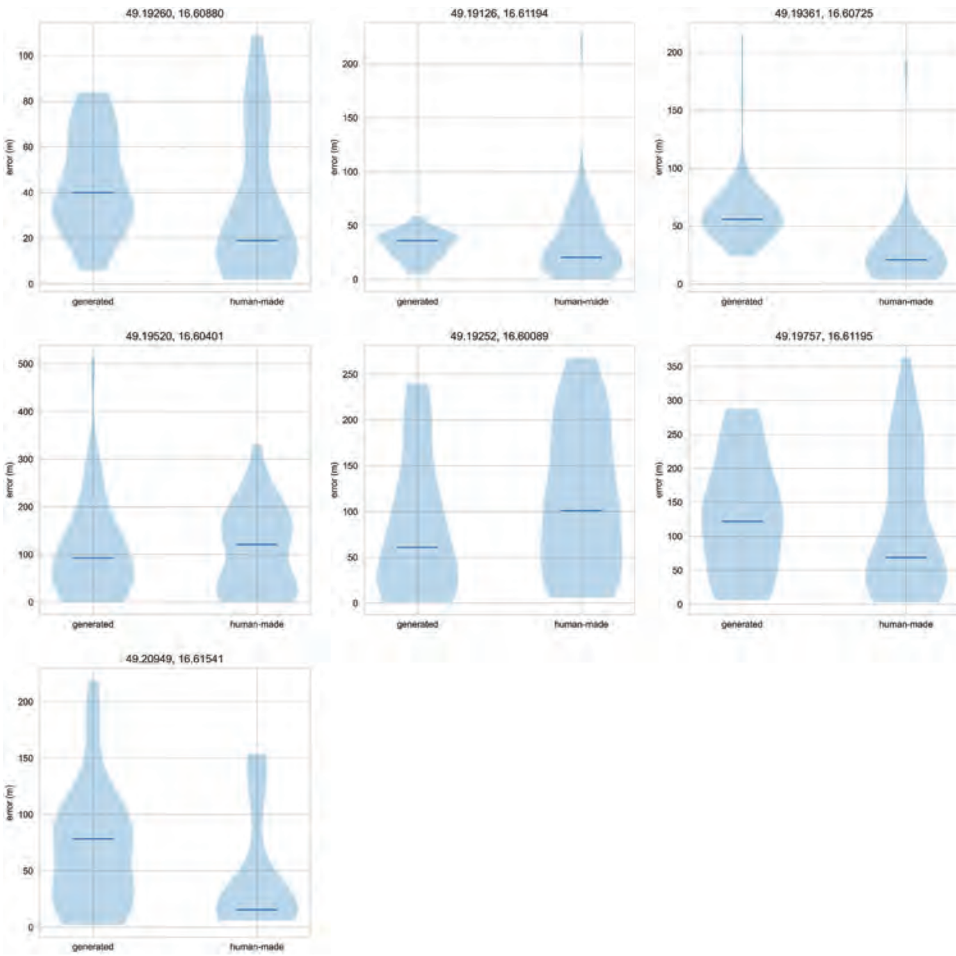

Figure 9. Errors in place identification using place descriptions – breakdown for each location.

For certain locations, the identification accuracies for human-made and generated descriptions were very similar. Such a situation arose when the method succeeded in capturing the relevant ROs and presented them in a understandable way. However, in certain cases, the proposed method may not have adequately assessed important ROs, leading some users to misidentify the place in question due to the suboptimal description.

We conducted an analysis to determine whether place descriptions generated by our method and human-made place descriptions were comparable in terms of identification accuracy. Following this, error rates were compared using unpaired two one-sided t-tests (TOST) introduced by Schuirmann (1981) with the following null hypothesis: The difference in error rate (m) between human-made and generated place descriptions is greater than or equal to 20 m. The null hypothesis was rejected with a p-value of 0.014. This

Figure 10. Results for the selected locations: map showing the true location and the assumed location.

suggests that there is no statistically significant difference between the types of descriptions in the 20 m similarity region (epsilon). It is important to note that the choice of epsilon affects the power of the result and was chosen rather conservatively.

Overall, the results suggest that the method manages to fulfill its intended objective of generating a place description in a natural language that leads to the correct identification of a place in a given context. Subjects in the test group were able to identify the place with an accuracy that is in general comparable to using human-made descriptions given the intended geographic scale, use cases, and scope. Despite the differences in identification accuracy, the proposed method is demonstrating the potential to generate place descriptions that facilitate place identification.

5.4. Limitations

When generalizing our results, it is important to acknowledge certain limitations. Specifically, our findings should be considered applicable to urban environments similar to those studied. This implies that the external validity of our findings may extend only to cities that share comparable characteristics with the studied area.

The application of the proposed method to other environments, such as rural areas, is in principle feasible. However, it would be necessary to assess the feature relevance for the target environment. Such adaptability could be the focus of further study.

In addition, the proposed method for generating place descriptions relies heavily on the availability of high-quality OpenStreetMap (OSM) data. In the absence of such data, the approach presented in this paper cannot be implemented.

In future work, it would be valuable to explore an end-to-end machine learning approach that would allow LLMs to learn important features of place descriptions, such as the number of ROs and their types. In addition, incorporating the selection of spatial relationships through a learned model such as Malinowski and Fritz (2013), rather than rule-based methods, could further enhance the capabilities of the proposed methods. However, a major challenge remains the lack of sufficiently large and diverse datasets for this specific task needed to support these advanced methods.

6. Conclusions

In this research, we investigated the possibility of using machine-generated descriptions to characterize a place so that a human can accurately identify it. We looked at how people use natural language to convey spatial information about a place. Based on our findings, we present a method for automatically generating place descriptions.

We devised a process that encompasses spatial data acquisition, relevance assessment, identification of spatial relationships, and generation of natural language descriptions. Our method combines rule-based generation of spatial relation facts, derived from OpenStreetMap data, with a LLM that converts these facts into natural language descriptions descriptions. In an experimental evaluation, we compared machine-generated place descriptions with human-made descriptions in a place identification task.

Overall, our results suggest that the proposed method is capable of describing a place in a way that is sufficient for humans to find it. The method consists of a rule-based process informed by a study of how people describe places, coupled with LLM. Such a technique could be used in various applications, such as navigation systems, virtual assistants, or location-based services. Automatically generated place descriptions can provide a more intuitive and accessible way to convey spatial information, especially in situations where visual media cannot be used, such as when interacting with devices without displays. Our research contributes to the growing body of knowledge on place descriptions and demonstrates the potential of an automated process to describe a place.

Disclosure statement

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

References

Anderson, J. (2016). *Editing density map*. MapBox. Retrieved July 27, 2023, from, http://mapbox.github.io/osm-analysis-collab/editor-density

Arora, S., Narayan, A., Chen, M. F., Orr, L., Guha, N., Bhatia, K., Chami, I., & Re, C. (2022). Ask me anything: A simple strategy for prompting language models. *The Eleventh International Conference on Learning Representations*, Kigali, Rwanda.

Arsanjani, J., Zipf, A., Mooney, P., & Helbich, M. (2015). *OpenStreetMap in GIScience: Experiences, Research and Applications*. Springer.

Axelsson, A., & Skantze, G. (2023). Using large language models for zero-shot natural language generation from knowledge graphs. *Proceedings of the Workshop on Multimodal, Multilingual Natural Language Generation and Multilingual WebNLG Challenge*. Association for Computational Linguistics, Prague, Czech Republic (pp. 39–54).

Balla, D., Zichar, M., Toth, R., Kiss, E., Karancsi, G., & Mester, T. (2020). Geovisualization techniques of spatial environmental data using different visualization tools. *Applied Sciences*, *10*(19), 6701. https://doi.org/10.3390/app10196701

Bezirhan, U., & Davier, M. (2023). Automated reading passage generation with openai's large language model. *Computers and Education: Artificial Intelligence*, *5*, 100161. https://doi.org/10.1016/j.caeai.2023.100161

Bhavya, B., Xiong, J., & Zhai, C. (2022, July). Analogy generation by prompting large language models: A case study of InstructGPT. *Proceedings of the 15th International Conference on Natural Language Generation* (pp. 298–312). Association for Computational Linguistics, Waterville, Maine, USA and virtual meeting.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., Hesse, C., & Chen, M. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901. https://doi.org/10.5555/3495724.3495883

Chung, J., Kamar, E., & Amershi, S. (2023). Increasing diversity while maintaining accuracy: Text data generation with large language models and human interventions. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, Toronto, Canada (pp. 575–593).

Clark, E., August, T., Serrano, S., Haduong, N., Gururangan, S., & Smith, N. A. (2021). All that's 'human'is not gold: Evaluating human evaluation of generated text. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, Bangkok, Thailand (pp. 7282–7296).

Cook, A., & Karakus, O. (2024). Llm-commentator: Novel fine-tuning strategies of large language models for automatic commentary generation using football event data. *Knowledge-Based Systems*, *300*, 112219. https://doi.org/10.1016/j.knosys.2024.112219

Davis, J. R., & Schmandt, C. M. (1989). The back seat driver: Real time spoken driving instructions. *Conference Record of Papers Presented at the First Vehicle Navigation and Information Systems Conference*, Toronto, Canada (pp. 146–150). Ieee.

De Carolis, B., Cozzolongo, G., Pizzutilo, S., & Silvestri, V. (2007). Mymap: Generating personalized tourist descriptions. *Applied Intelligence*, *26*(2), 111–124. https://doi.org/10.1007/s10489-006-0012-4

De Carolis, B., Novielli, N., Plantamura, V. L., & Gentile, E. (2009). Generating comparative descriptions of places of interest in the tourism domain. *Proceedings of the Third ACM Conference on Recommender Systems*, New York, USA (pp. 277–280).

Dethlefs, N., Wu, Y., Kazerani, A., & Winter, S. (2011). Generation of adaptive route descriptions in urban environments. *Spatial Cognition & Computation: An Interdisciplinary Journal*, *11*(2), 153–177. https://doi.org/10.1080/13875868.2010.525272

Dzial-Owski, K. (2021). *OSM scout server. GitHub repository*. Retrieved July 3, 2023, from https://github.com/rinigus/osmscout-server

Gao, T., Fisch, A., & Chen, D. (2021). Making pre-trained language models better few-shot learners. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing* (pp. 3816–3830).

Garfinkel, H. (1967). *Studies in ethnomethodology*. Prentice-Hall.

Hakobyan, L., Lumsden, J., O'Sullivan, D., & Bartlett, H. (2013). Mobile assistive technologies for the visually impaired. *Survey of Ophthalmology*, *58*(6), 513–528. https://doi.org/10.1016/j.survophthal.2012.10.004

Hall, M. M., & Jones, C. B. (2022). Generating geographical location descriptions with spatial templates: A salient toponym driven approach. *International Journal of Geographical Information Science*, *36*(1), 55–85. https://doi.org/10.1080/13658816.2021.1913498

Hall, M. M., Jones, C. B., & Smart, P. (2015). Spatial natural language generation for location description in photo captions. *Spatial Information Theory: 12th International Conference, COSIT 2015*, Santa Fe, USA (pp. 196–223). Springer.

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

Harkous, H., Groves, I., & Saffari, A. (2020). Have your text and use it too! end-to-end neural data-to-text generation with semantic fidelity. *Proceedings of the 28th International Conference on Computational Linguistics*. International Committee on Computational Linguistics, Barcelona, Spain (pp. 2410–2424).

Herskovits, A. (1985). Semantics and pragmatics of locative expressions. *Cognitive Science*, *9*(3), 341–378. https://doi.org/10.1016/S0364-0213(85)80003-3

Jin, F., Lu, J., Zhang, J., & Zong, C. (2023). Instance-aware prompt learning for language understanding and generation. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *22*(7), 1–18. https://doi.org/10.1145/3604613

Karpinska, M., Akoury, N., & Iyyer, M. (2021). The perils of using mechanical Turk to evaluate open-ended text generation. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Punta Cana, Dominican Republic (pp. 1265–1285).

Kasner, Z., Garanina, E., Platek, O., & Dusek, O. (2023). Tabgenie: A toolkit for table-to-text generation. *Proceedings of 61st Annual Meeting of the Association for Computational Linguistics*, Toronto, Canada (pp. 444–455). Association for Computational Linguistics.

Keymanesh, M., Benton, A., & Dredze, M. (2022). What makes data-to-text generation hard for pretrained language models? *Proceedings of the 2nd Workshop on Natural Language Generation, Evaluation, and Metrics*, Abu Dhani, UAE (pp. 539–554). Association for Computational Linguistics.

Klopfenstein, L. C., Delpriori, S., Malatini, S., & Bogliolo, A. (2017). The rise of bots: A survey of conversational interfaces, patterns, and paradigms. *Proceedings of the 2017 conference on designing interactive systems*, Edinburgh, United Kingdom (pp. 555–565).

Kuipers, B. (1978). Modeling spatial knowledge. *Cognitive Science*, *2*(2), 129–153. https://doi.org/10.1016/S0364-0213(78)80003-2

Landau, B., & Jackendoff, R. (1993). "What" and "where" in spatial language and spatial cognition. *Behavioral and Brain Sciences*, *16*(2), 217–238. https://doi.org/10.1017/S0140525X00029733

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity*. Cambridge University Press.

Li, J., Tang, T., Zhao, W. X., & Wen, J. R. (2022). Pretrained language models for text generation: A survey. *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, Montreal, Canada (pp. 4492–4499). International Joint Conferences on Artificial Intelligence.

Lin, Y., Ruan, T., Liu, J., & Wang, H. (2023). A survey on neural data-to-text generation. *IEEE Transactions on Knowledge and Data Engineering*.

Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, *55*(9), 1–35. https://doi.org/10.1145/3560815

Malinowski, M., & Fritz, M. (2013). Learning smooth pooling regions for visual recognition electronic. In T. Burghardt, D. Damen, W. Mayol-Cuevas & M. Mirmehdi (Eds.), Proceedings of the British Machine Vision Conference Durham (pp. 1–11). England.

Michelet, G., & Breitinger, F. (2024). Chatgpt, llama, can you write my report? an experiment on assisted digital forensics reports written using (local) large language models. *Forensic Science International: Digital Investigation*, *48*, 301683. https://doi.org/10.1016/j.fsidi.2023.301683

Min, S., Lyu, X., Holtzman, A., Artetxe, M., Lewis, M., Hajishirzi, H., & Zettlemoyer, L. (2022). Rethinking the role of demonstrations: What makes in-context learning work? *Proceedings* *of the 2022 Conference on Empirical Methods in Natural Language Processing*, Abu Dhabi, UAE (pp. 11048–11064).

Minghini, M., & Frassinelli, F. (2019). OpenStreetMap history for intrinsic quality assessment: Is osm up-to-date? *Open Geospatial Data, Software and Standards*, *4*(1), 1–17. https://doi.org/10.1186/s40965-019-0067-x

Moiseev, F., Dong, Z., Alfonseca, E., & Jaggi, M. (2022). SKILL: Structured knowledge infusion for large language models. *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Seattle, USA (pp. 1581–1588). Association for Computational Linguistics.

Mosbach, M., Pimentel, T., Ravfogel, S., Klakow, D., & Elazar, Y. (2023). Few-shot fine-tuning vs. in-context learning: A fair comparison and evaluation. In A. Rogers, J. Boyd-Graber, & N. Okazaki (Eds.), *Findings of the association for computational linguistics: ACL 2023* (pp. 12284–12314). Association for Computational Linguistics.

Naaman, M., Song, Y. J., Paepcke, A., & Garcia-Molina, H. (2004). Automatic organization for digital photographs with geographic coordinates. *Proceedings of the 4th ACM/IEEE-CS joint conference on Digital libraries*, Tuscon, USA (pp. 53–62).

Nan, L., Radev, D., Zhang, R., Rau, A., Sivaprasad, A., Hsieh, C., Tang, X., Vyas, A., Verma, N., Krishna, P., Liu, Y., Irwanto, N., & Pan, J. (2021). DART: Open-domain structured data record to text generation. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, Mexico City, Mexico (pp. 432–447).

Oliveira, V. M. A. D. (2021). *Urban morphology: An introduction to the study of the physical form of cities* (2nd ed.). Springer Cham.

OpenAI. (2023a). *Fine-tuning*. Retrieved August 10, 2023, from https://beta.openai.com/docs/guides/fine-tuning/

OpenAI. (2023b). *Openai platform reference*. Retrieved August 10, 2023, from https://platform.openai.com/docs/api-reference

OSM Contributors. (2023). *Map features*. Retrieved July 27, 2023, from, https://wiki.openstreetmap.org/wiki/Map_features

Raaflaub, K. A., & Talbert, R. J. (2009). *Geography and ethnography: Perceptions of the world in pre-modern societies*. John Wiley & Sons.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, *1*(8), 9.

Reiter, E., Sripada, S., Hunter, J., Yu, J., & Davy, I. (2005). Choosing words in computer-generated weather forecasts. *Artificial Intelligence*, *167*(1–2), 137–169. https://doi.org/10.1016/j.artint.2005.06.006

Richter, D., Vasardani, M., Stirlng, L., Richter, K. F., & Winter, S. (2013). Zooming In–Zooming Out Hierarchies in Place Descriptions. *Progress in Location-Based Services* (pp. 339–355). Springer. https://doi.org/10.1007/978-3-642-34203-5_19

Richter, D., Winter, S., Richter, K. F., & Stirling, L. (2012). How people describe their place: Identifying predominant types of place descriptions. *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information*, Redondo Beach, USA (pp. 30–37).

Richter, D., Winter, S., Richter, K. F., & Stirling, L. (2013). Granularity of locations referred to by place descriptions. *Computers, Environment and Urban Systems*, *41*, 88–99. https://doi.org/10.1016/j.compenvurbsys.2013.03.005

Scao, T. L., Fan, A., Akiki, C., Pavlick, E., Ilic, S., Hesslow, D., Castagn´e, R., Luccioni, A. S., Yvon, F., Gall´e, M., Tow, J., Rush, A. M., Biderman, S., Webson, A., Ammanamanchi, P. S., Wang, T., Sagot, B., Muennighoff, N. . . . Adelani, D. I. (2022), preprint. BLOOM: A 176B-Parameter open-access multilingual language model. *arXiv preprint arXiv: 2211.05100*, abs/2211.05100.

Schuirmann, D. (1981). On hypothesis testing to determine if the mean of a normal distribution is contained in a known interval. *Biometrics*, *37*(3), 617–617.

Stock, K., & Yousaf, J. (2018). Context-aware automated interpretation of elaborate natural language descriptions of location through learning from empirical data. *International Journal of Geographical Information Science*, *32*(6), 1087–1116. https://doi.org/10.1080/13658816.2018.1432861

Turner, R., Sripada, S., & Reiter, E. (2009). Generating approximate geographic descriptions. *Proceedings of the 12th European Workshop on Natural Language Generation*, Athens, Greece (pp. 121–140). Springer.

Walmsley, D. (1974). Emotional involvement and subjective distance: A modification of the inverse square root law. *The Journal of Psychology*, *87*(1), 9–19. https://doi.org/10.1080/00223980.1974.9915668

Wang, Q., Pan, X., Huang, L., Zhang, B., Jiang, Z., Ji, H., & Knight, K. (2018). Describing a knowledge base. *Proceedings of the 11th International Conference on Natural Language Generation*. Association for Computational Linguistics, Tilburg, Netherlands (pp. 10–21).

Weilenmann, A. H., & Leuchovius, P. (2004). I'm waiting where we met last time exploring everyday positioning practices to inform design. *Proceedings of the Third Nordic Conference on Human-Computer Interaction*, Tampere, Finland (pp. 33–42).

Winter, S., Richter, K. F., Baldwin, T., Cavedon, L., Stirling, L., Duckham, M., Kealy, A., & Rajabifard, A. (2011). Location-based mobile games for spatial knowledge acquisition. Proceedings of the Workshop on Cognitive Engineering for Mobile GIS (COSIT 2011), Belfast, Maine, USA, *780*, 1–8.

Winter, S., & Truelove, M. (2013). Talking About Place Where it. In M. Matters Raubal, D. Mark, & A. Frank (Eds.), *Cognitive and Linguistic Aspects of Geographic Space: New Perspectives on Geographic Information Research* (pp. 121–139). Springer. https://doi.org/10.1007/978-3-642-34359-9_7

Zhao, Y., Zhang, H., Si, S., Nan, L., Tang, X., & Cohan, A. (2023). Investigating table-to-text generation capabilities of large language models in real-world information seeking scenarios. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track*, Singapore (pp. 160–175). Association for Computational Linguistics.

Zhou, C., Ludford, P., Frankowski, D., & Terveen, L. (2005). Talking about place: An experiment in how people describe places. *Proceedings of the Third International Conference on Pervasive Computing*, Munich, Germany (Vol. 3468. pp. Xiii, 347). Springer, Lecture Notes in Computer Science.

Appendix

Place description used for evaluation

Place: 49.19260, 16.60880

- *Machine-generated*: You are located on Radnická Street at Zelný Trh. In the immediate vicinity you'll find the Bucheck grocery store with the Brno Dragon artwork behind it. A little further is the Tržnice market and the Reduta Theatre. The Dietrichstein Palace is also nearby.
- *Human-made*: It's a place on the square called Zelný trh, located in the very center of the city. The square is approximately square in shape with a baroque fountain in the center. On the northern side of the square there is a modern department store, and next to it, on the corner of Radniční Street, is La Formaggeria Gran Moravia, a cheese shop. The place you are looking for is on the sidewalk in between the department store and the cheese shop.
- *Human-made*: The place is in the middle of Zelný trh, behind the place you are looking for is Bucheck.
- *Human-made*: Near the fountain at Zelný trh. The place is a little to the left of La Formaggeria Gran Moravia (cheese shop), but before the greengrocer's market.

Place: 49.19126, 16.61194

- *Machine-generated*: This place is in the city center on Nádražní Street. It is in front of the Hlavní nádraží public transport stop and the KFC fast food restaurant. The Letmo shopping center with the Albert supermarket is also nearby.
- *Human-made*: It's next to the KFC at the main train station, at the farthest public transport stop from the train station, the one closest to the city center.
- *Human-made*: The place is near the tram stop in front of the main train station, more precisely the platform by the KFC and Kebab shop.
- *Human-made*: The square in front of the main train station in Brno, on the tram tracks, the last track from the station building.

Place: 49.19361, 16.60725

- *Machine-generated*: You are on the Dominican Square next to the Church of St. Michael the Archangel. Behind the church, but a little further away, there is the Velký Špalíček shopping center. In Velký Špalíček there is the Tesco supermarket and a movie theater. A little further away is the House of the Lords of Kunstat.
- *Human-made*: It's on the street across from St. Michael's Church (a large cathedral) and right next to it, there is the Burkalo shop.
- *Human-made*: Near the main entrance to Velký Špalíček, in the direction of Česká Street, in front of St. Michael's Church.
- *Human-made*: The place is near the front of the department store Velký Špalíček, at the end of Veselá Street, in front of St. Michael's Church.

Place: 49.19520, 16.60401

- *Machine-generated*: This place is located in the center of the city near Husova Street. Next to the street is Špilberk Park and tram line 5. The Beseda House is not far away. A little further away is the Museum of Applied Arts.
- *Human-made*: Beneath Špilberk Castle, near the small entrance at the crossroads across from the hotel, there is a red church at the left end of the street and a bus stop and Petrov at the right end.

- *Human-made*: The place is on the edge of a city park, opposite a parking lot and a parking garage. You are on the side of the street opposite the parking lot.
- *Human-made*: The place is on Husova Street, facing the main road. Across the street is a large parking lot belonging to a hotel. Behind you is M. Jakszta Park. In front of you, slightly to the left, is Pražákův Palace. If you walk along the street, you will come across a school.

Place: 49.19252, 16.60089

- *Machine-generated*: You are near Pellicova Street in the Stare Brno district, right next to Špilberk. U Alberta pub is located in the park. A little further there is the tram line 6 and the restaurant Shanghai. Further away is the Bethlehem Church.
- *Human-made*: The place is located on Pellicova Street, which runs under the southern slope of the hill on which Špilberk Castle stands. It is located south of the U Alberta pub, at the place where the buildings start, not far from the playground.
- *Human-made*: If you go down Pellicova Street, you'll find a pet food store, and the place is at the end of the path across the street from the park.
- *Human-made*: The place is on Pellicova Street near Hofa company.

Place: 49.19757, 16.61195

- *Machine-generated*: You are located in Janáček Park, right next to tram line 1. The tram line is in front of the Janáček Theatre. Nearby are the Palace of Justice and the Brno Regional Court. A little further is Roosevelt Street, and even further is KolištéPark.
- *Human-made*: The place is located behind the Janáček Theatre, heading from Moravia Square down to the intersection where the Music Club is.
- *Human-made*: At the back corner of the Janacek Theater. It is a small park, near the pedestrian crossing and parking lot.
- *Human-made*: It's on the back side of the Janáček Theatre, on the opposite side of the tram line where a liquor store is located.

Place: 49.20949, 16.61541

- *Machine-generated*: This place is on Zeměědělská Street near the campus of Mendel University in Brno. Nearby, you'll find the Erbenova public transport stop and a parking lot. Na chate restaurant is a short distance away, with trees behind it.
- *Human-made*: The entrance to the Faculty of Business on Zemědělská Street, a short walk from the Erbenova bus stop.
- *Human-made*: The entrance to the Mendel University campus, at the bus stop – Erbenova.
- *Human-made*: The place is near the gate in front of the Q building of Mendel University in Brno.

Comparison to Google Bard with Maps extension enabled

Highlighted parts are not accurate.

Place at 16.6152, 49.1831

Is a brownfield with few notable features, the most prominent being the sparsely used railroad station.

Google Bard with Google Maps

This place is a street called Rosicka. It is located in the Brno-střed district of Brno, Czech Republic. The street is approximately **0.2 kilometers** long and runs east-west. It is a **residential street** with a mix of houses and apartments. There are also a **few shops and restaurants** on the street. Here are some additional details about the place: It is a **quiet and peaceful street** with a **relaxed atmosphere**. It is a good place to live if you are looking for a convenient location with easy access to amenities. It is also a good place to visit if you are interested in seeing a **typical Brno neighborhood**.

Proposed method

You are located in the Trnitá district on Rosická Street at the bus stop Dolní nádraží. On the same street, you are close to a railway track and a bit further there is a park and a parking. In a greater distance, there is the railway station Brno dolní nádraží.