

User-Driven Toponym Disambiguation Using Dialogue

Mikuláš Muroň, František Dařena, David Procházka & Roman Kern

To cite this article: Mikuláš Muroň, František Dařena, David Procházka & Roman Kern (2023) User-Driven Toponym Disambiguation Using Dialogue, Journal of Map & Geography Libraries, 19:3, 198-219, DOI: [10.1080/15420353.2024.2363212](https://doi.org/10.1080/15420353.2024.2363212)

To link to this article: <https://doi.org/10.1080/15420353.2024.2363212>



© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 17 Jun 2024.



Submit your article to this journal [↗](#)



Article views: 101





View related articles [↗](#)



View Crossmark data [↗](#)

User-Driven Toponym Disambiguation Using Dialogue

Mikuláš Muroň^a , František Dařena^a, David Procházka^a  and Roman Kern^{b*} 

^aDepartment of Informatics, Faculty of Business and Economics, Mendel University in Brno, Brno, Czech Republic; ^bKnow-Center, Research Center for Data-Driven Business & Big Data Analytics, Graz, Austria

ABSTRACT

The paper presents a novel method for processing information with a spatial component or querying geospatial databases. It proposes an interactive toponym disambiguation method tailored especially for dialogue systems and chatbots. The method exploits the interactive nature of dialogues to resolve ambiguity by using dialogue clarification techniques, i.e., asking users for additional information to perform disambiguation. The paper evaluates different questioning strategies and methods for selecting suitable toponym features to formulate disambiguation questions. A novel Equi-population partitioning method for selecting suitable toponym features in questions is presented as an approach outperforming other compared methods.

ARTICLE HISTORY

Received 11 January 2024

Revised 28 March 2024

Accepted 30 May 2024

KEYWORDS

toponym disambiguation;
dialog systems; dialog
clarification

INTRODUCTION

One of the many roles of libraries is to facilitate access to information, acting as intermediaries and providers of knowledge (Evans and Baker 2011; Chakrabarti and Mahapatra 1989). Historically, information has been available in a physical form, such as a paper catalog. Technological advances have not only digitized this repository of information, but also introduced new and previously unimagined ways of interacting with it, such as criterion search, full-text search, and map search. However, current methods of searching for information with a spatial component are not the final frontier.

In recent years, we have seen a notable shift toward more interactive forms of human-computer interfaces: chatbots and conversational interfaces. In certain scenarios, the use of natural language provides a more intuitive way to interact with computers, as evidenced by the growing popularity of such interfaces (Hadi Mogavi et al. 2024; Garrel and Mayer 2023). Conversational interfaces are changing the way we inquire and obtain

CONTACT Mikuláš Muroň  mikulas.muron@gmail.com  Department of Informatics, Mendel University in Brno, Brno, Czech Republic.

*Additional affiliation: Know-Center GmbH & Institute of Interactive Systems and Data Science Graz University of Technology, Graz, Austria.

© 2024 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

information in various domains (Adamopoulou and Moussiades 2020; Fu, Mouakket, and Sun 2023; Suhaili, Salin, and Jambli 2021), and have the potential to serve as a tool for querying and retrieving information.

Consequently, chatbots and interactive dialog systems in general are increasingly finding their niche in the library context (Sanji, Behzadi, and Gomroki 2022; Rodriguez and Mune 2022; Reinsfelder and O'Hara-Krebs 2023; DeeAnn 2012), being tested and even utilized in a wide range of applications, including information and reference services (Ehrenpreis and DeLooper 2022; Nawaz and Saldeen 2020; Panda and Chakravarty 2022; McNeal and Newyear 2013; Vincze 2017). Information retrieval could be another promising application domain for chatbots in a library context, allowing such systems to effectively query catalogs or geospatial databases. Non-automated conversational interfaces have begun to help people find the right spatial data at the GIS section of the University of North Carolina library (Scaramozzino et al. 2014), and there has been research into facilitating natural language interaction with geospatial databases through projects like GeoDialogue (Cai et al. 2005). Modern systems such as Lafia et al. (2019) or ESRI (2018) go beyond search, and aim to perform some of the GIS tasks using a chatbot. In such scenarios, a system needs to understand and interpret natural language queries correctly to provide accurate answers.

The necessary steps for understanding spatial references in user queries are toponym recognition (i.e., identification of a place name in an unstructured text) and toponym disambiguation (i.e., assigning an ambiguous place name to one particular reference). This whole process is referred to as geoparsing (Gritta, Pilehvar, and Collier 2020). Finding a solution to this problem would help spatial query resolution, which could be applied not only in the context of library spatial query search but in any spatially focused task.

Problem Statement

Geographic libraries are essential for collecting, organizing, and providing access to geographic information. The ability to track and retrieve locations associated with information resources is important for documenting both provenance and subject matter (Radio et al. 2021). When searching a map, geospatial dataset, or other resource with a spatial component, the challenge of disambiguation arises as a queried place name may refer to several possible candidates. Algorithms enabling automated identification of named places and accurate place disambiguation are thus needed (Brando and Frontini 2017). Candidates are often contained in structured knowledge sources (such as Geonames.org), to which libraries' geographic datasets are often connected (Butler et al. 2017). While there are several methods for place name disambiguation, none of them focus on disambiguation

tailored to dialog systems. Interactive dialog interfaces have several specifics that should be reflected in the disambiguation method.

In this paper, we focus on the second part of the geoparsing process, toponym disambiguation. The disambiguation task can be illustrated as follows: Given the list of toponyms extracted from the natural language query provided by the user (e.g., “Find me all maps of San Antonio” → San Antonio), the task is to select the candidate intended by the user, since the toponym itself (“San Antonio”) is ambiguous and can refer to multiple places with the same name. Compared to the existing disambiguation methods that are mostly based on heuristics (simple rules to solve complex problems by making practical shortcuts), we aim to exploit the interactive nature of a conversation to resolve toponym ambiguity by deploying dialog clarification strategies. As shown by Braslavski et al. (2017), this approach has many advantages. We aim to resolve the ambiguity by asking the user to narrow down a set of possible geographical locations. Therefore, the dialog itself serves as a source of additional information. To the best of our knowledge, no work addresses the problem of toponym disambiguation in dialog systems.

Research Objective

Although various methods for toponym disambiguation currently exist (Kafando et al. 2023; Rauch, Bukatin, and Baker 2003; Leidner 2007; Andogah 2011), their use in the domain of dialog systems/chatbots is not well-explored. Moreover, the existing methods are suited for cases when enough context (e.g., a longer text, the general location is known, etc.) is provided (Adamopoulou and Moussiades 2020; Kafando et al. 2023). The challenge arises from the different nature of dialogs which, compared to traditional unstructured texts like news or academic papers, are characterized by their brevity, featuring short sentences often devoid of the contextual cues typically found in longer texts. For instance, in a travel blog, the author might discuss Paris, referencing its landmarks, thereby providing context to ascertain that Paris in France is intended. Conversely, in a chatbot interaction, a user might simply ask, “What’s the weather in Paris?” without any additional context.

This research aims to fill the existing gap in automated place name disambiguation methods by developing and evaluating a method that can accurately resolve toponym ambiguity in the context of conversational interfaces. The principal goal is to improve the functionality and reliability of chatbots and interactive dialogue systems in libraries and other information-centric contexts, enabling them to correctly interpret spatial user queries and thus serve as tools for querying and retrieving resources with a spatial component.

CURRENT STATE

Geoparsing is a process of converting unstructured text containing place names (toponyms) into a set of unambiguous identifiers (geographic coordinates or references to unique entities in a gazetteer or knowledge base). Geoparsing is composed of two distinct steps: geotagging and geocoding. The main goal of geotagging (in literature also known as geographical entity extraction) is to extract place names (toponyms) from ingested unstructured text. The second step in the geoparsing process is geocoding. It aims to perform toponym disambiguation (a specialized variant of named entity disambiguation) and consequently entity linking. Toponym disambiguation is a task that aims to correctly assign an ambiguous place name to a specific reference (Buscaldi and Rosso 2008). The general scheme of a geoparsing system is shown in Figure 1.

In this paper, we focus on the second step of the geoparsing pipeline, i.e., geocoding, also known as toponym disambiguation. We are purposely omitting the first part of the pipeline; the geotagging step (a subset of named entity recognition, NER), as there has been tremendous progress in the last few years with the introduction of transformers and large pre-trained language models (Zhou et al. 2020; Khan et al. 2023).

The knowledge base or gazetteer is an essential part of the toponym disambiguation process. Knowledge bases and gazetteers have proven to be suitable resources, especially when the messages are short and informal (Habib and Van Keulen 2016). Radio et al. (2021) conducted an assessment of geographic vocabulary usage and while many various datasets exist (e.g., Wikidata, Geo-WordNet, and Open Street Maps), many researchers (Santos et al. 2018; El Midaoui et al. 2018; Acheson, Volpi, and Purves 2020) and even commercial projects rely on the Geonames gazetteer (<http://www.geonames.org>) as the main data source for toponym extraction and entity linking (Kejriwal and Szekely 2017). This can be attributed to the fact that Geonames is one of the most comprehensive gazetteers currently available. It also provides not only place names and their spatial

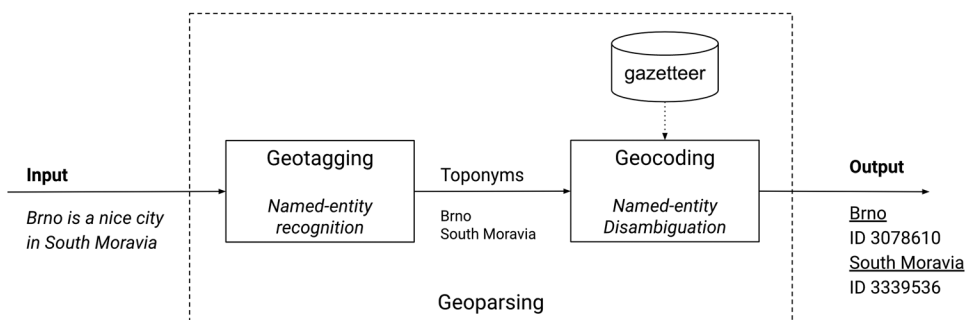


Figure 1. Scheme of a general geoparsing system consisting of the geotagging and geocoding stages proposed by Gritta, Pilehvar, and Collier (2020).

component but also additional features (place type and class, country, and administrative region, etc.) for each entry. It is distributed under the Creative Commons license.

It is relevant to know how often toponym ambiguity occurs to perceive the scope of the problem. Our analysis, based on the Geonames gazetteer containing 11.8 million toponyms, involved marking each entry for the presence of any homonyms. This process revealed that approximately 16% of all entries correspond to more than one location. [Figure 2](#) depicts the distribution of ambiguous toponyms in the Geonames dataset, i.e., the relationship between the number of possible candidates and their occurrence. Although the actual prevalence of ambiguous toponyms depends on the use case, the size of the area (i.e., we can consider street names to be unambiguous within a city, but not within a country), and domain specifics, it occurs to some extent in most scenarios (see [Figure 3](#) for illustration) and is therefore a problem worth considering.

Toponym Disambiguation Methods

We can consider the toponym disambiguation problem to be a specific case of the more general problem of Word Sense Disambiguation (WSD). Although a variety of WSD methods and specialized toponym disambiguation methods are described in the literature, some issues still linger. Most existing methods rely on some kind of heuristic to estimate the right candidate. The term heuristic in this context describes a method that provides an educated guess to the problem, based on practical experience or judgment.

To disambiguate a toponym, information about context toponyms (i.e., toponyms contained in the same document) is often used. Other external information that can be considered is, e.g., social media tags (Middleton

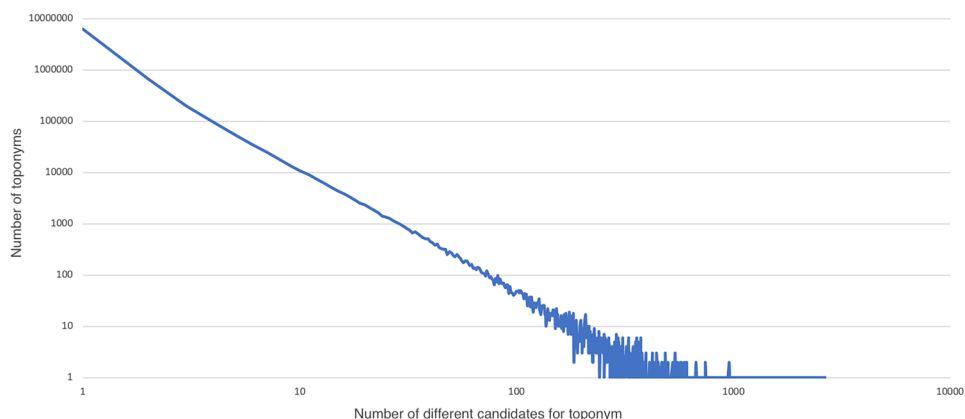


Figure 2. The distribution of ambiguous toponyms in the Geonames gazetteer. The relationship between the number of possible candidates and their occurrence resembles the power-law relationships.

Nova Scotia

Italian tourists end up in wrong Sydney

Travel agency's booking error landed couple in Nova Scotia instead of Australia

CBC News · Posted: Jul 07, 2010 4:38 PM EDT | Last Updated: July 8, 2010

A couple from Italy got a taste of Cape Breton hospitality Wednesday when they unexpectedly arrived in Sydney, N.S., instead of Sydney, Australia.

Figure 3. Excerpt from CBC News: *Italian tourists end up in wrong Sydney* (CBC News 2010).

et al. 2018), a human-made list of words that form collocations with toponyms (Boyarsky, Kanevsky, and Butorina 2020), information about toponyms of connected users in a social network (Ghufran, Quercini, and Bennacer 2015), or the affiliation of the author of a text (Torvik 2015).

The context toponyms can be the input to an approach relying on the spatial minimality heuristics proposed by Leidner (2007). It assumes that toponyms in a document tend to spatially cluster together. In case of ambiguity, we should therefore pick the candidate, which minimizes the distance to the center of the cluster of possible candidates. Other heuristics work with, for example, spatial proximity (Andogah 2011) or spatial hierarchy (Karimzadeh et al. 2013).

The existing methods can perform well if they are applied in cases where enough context is provided. This is typically the case when the system processes longer texts containing many different toponyms so that the context can be inferred. Problems can occur when working with short texts, texts with nonstandard syntax, or abbreviations, which is typical, e.g., for microblog posts (Karimzadeh et al. 2013). Similar challenges can be expected in dialogs because many characteristics of short microblog posts are in some aspects similar to dialogs.

Other heuristics must be thus used when the information about the context according to which a toponym can be disambiguated is missing or is not satisfactory. One of the simplest used heuristics is the population heuristics discussed by Rauch, Bukatin, and Baker (2003), who suppose that highly populated places tend to be mentioned more frequently than places with low populations. This approach is in some aspects similar to the Most Frequent Sense heuristic (Calvo and Gelbukh 2014), which assumes that the most common meaning of a word is the correct one. For instance, the word 'bank' would be always interpreted as a financial institution rather than the side of a river, based on its more frequent

usage in that sense. Although this method has proven to be a competitive baseline in general WSD, it has many drawbacks, especially poor performance in the recall metric, as discussed by Lieberman and Samet (2012). In response to these challenges, we aim to take a different approach and leverage the interactive nature of a conversation to resolve a toponym's ambiguity. We propose to engage the user to assist the system in the toponym disambiguation process by means of a clarification dialog. Clarification dialogs, which involve requesting additional information or clarification from the conversation partner to resolve ambiguities or misunderstandings, are frequently employed in human conversations, occurring in 3%–6% of dialogue turns (Purver, Ginzburg, and Healey 2003). Not surprisingly, strategies mimicking this behavior have been studied for utilization in dialog systems and chatbots (Codem et al. 2015). Inspired by this stream of research, we aim to adapt the clarification strategies proposed in the previously published literature for the task of toponym disambiguation.

Using Dialogue for Disambiguation

To perform interactive toponym disambiguation, we propose utilizing established dialog clarification strategies suited for cases where: (i) a finite set of possible candidates exists, and (ii) properties describing the ambiguous candidates are available.

We follow a straightforward dialogue clarification approach known as the filtering strategy (Soo and Cheng 2002). It is based on reducing the number of possible candidates by defining a sequence of conditions that enable sequential filtering of the candidates. It must be determined which feature of the candidate should be selected to compose a question and what value of the feature should be used in the question. When the properties are known (e.g., extracted from a gazetteer) they can be used to formulate a clarification question (condition) automatically, for example: *Do you mean San Antonio in the USA?*

It is necessary to consider two aspects—the first one, which is related to the information need of a user (e.g., a user wants San Antonio in the USA and not in Chile), and the other, which is related to the efficiency of the search. The information need cannot be recognized automatically without additional information and the user needs to be asked. A positive answer can filter out all candidates that are not in the USA.

The problem is in some aspects similar to the popular game Twenty Questions (Dagan et al. 2017), a guessing game where players try to identify an unknown entity in less than 20 questions. In the toponym disambiguation scenario, the system has a list of candidates, but only the user knows the correct one. The system can then use the features of the

candidates to ask questions (about the country, type of place, etc.). The most efficient approach is to formulate a question that can split possible candidates into two parts of almost equal size (Soo and Cheng 2002; Shannon 1948).

Approaches based on separating candidates into parts of the same size assume that all candidates are equally likely to occur. However, questions may lead to the correct answer, i.e., the identification of the right candidate, faster if the probability of some answers differs from others. We then can consider this problem from the perspective of information entropy. The final decision (disambiguating a certain toponym) can be made after asking several questions where every answer can bring us some amount of information. A measure that quantifies the average amount of information conveyed by an answer is known as information entropy (Shannon 1948). The highest number of questions on average needs to be asked (the messages have the highest entropy) when all the answers have the same probability. If the probability of possible answers is different, a lower number of questions needs to be asked on average (the messages have lower entropy) when all the answers have the same probability. If the probability of possible answers is different, a lower number of questions needs to be asked on average (the messages have lower entropy).

The assumption of different probabilities is generally useful for toponym disambiguation. People don't use toponyms with the same probability—in a given context some candidates have a higher prior probability to occur than others (e.g., tourists will more often buy an airplane ticket to Sydney, Australia than to Sydney, Canada as illustrated by Figure 3).

Suresh (2017) and Hu et al. (2018) argue that methods based on separating candidates into subsets with similar sizes have limitations that make them less suited for practical use. Although these methods are guaranteed to eventually find the right candidate (assuming that the candidates can be distinguished perfectly based on their feature values), their use may be impractical due to the possibly high number of questions asked. For illustration, one of the most common toponyms found in Geonames is *San Antonio*, which refers to 1,536 different places (Buscaldi 2010). As a consequence, an average of 10.58 questions would be needed to successfully disambiguate such a frequent toponym by applying an Equi-size partitioning approach (see later). Such a high number of questions would require truly devoted users, who might not represent the general public. Following the recommendation of Suresh (2017), who suggests constraining the number of questions to 5, we focus on finding a disambiguation strategy that would enable reaching the correct answer in fewer steps.

INTERACTIVE TOPONYM DISAMBIGUATION

We propose an interactive toponym disambiguation method that follows the principles of dialog clarification. We specifically target resolving the ambiguity of toponyms to perform better in real-world scenarios.

In these scenarios, users ask questions or request services that include names of geographic locations that can be ambiguous. No additional information (context) regarding the locations is generally available. It is required that the name is correctly assigned to one specific place. To distinguish the possible location candidates, their characteristics (usually contained in a source of structured information) need to be known so they can be used in the disambiguation process. The properties of candidates are used to formulate questions that, after being answered by a user, help eliminate the list of possible candidates. Besides the need for correct identification, the speed of disambiguation and simplicity of the interaction are emphasized.

The toponym disambiguation method addresses the problem of (i) formulating appropriate questions and (ii) determining their order. The appropriateness of questions and their order is assessed with respect to the speed of disambiguation.

To distinguish between toponyms, we need to have a sufficient number of features characterizing them, so questions can be formulated upon them. For this purpose, we used the attributes available from the Geonames gazetteer and selected those that are suitable for disambiguation. These are those for which we can expect that their value is known by users, and which capture a variety of aspects of the candidates. We ultimately arrived at this selection of features with an example of the data and feature values in [Table 1](#):

- Continent
- Country name
- Administrative region at all levels
- Feature code (e.g., farm village and capital of a political entity, etc.)
- Feature class (e.g., populated place, area, region, and vegetation, etc.)
- Time zone

Table 1. Example of features extracted from Geonames for a toponym *San Antonio* (only three out of 1,536 entries are shown).

ID	Continent	Country name	Administrative region (levels)	Feature Code	Feature Class	Population	Time zone
4726206	North America	United States	Texas, Bexar	seat of a second order administrative division	Populated place (city)	1 434 625	GMT-5
2410805	South America	Chile	Valparaíso	town, village, or other agglomeration	Populated place (city)	85 651	GMT-3
3621967	Central America	Costa Rica	Heredia, Belén	town, village, or other agglomeration	Populated place (city)	10 938	GMT-6

To ask fewer questions on average, we need additional information to determine the candidate's prior probability. In the case of toponym disambiguation, the heuristic relating the population of a candidate can be used (Rauch, Bukatin, and Baker 2003). This is based on the assumption that a place with a high population is more likely to be mentioned than a place with a lower population. Qi et al. (2019) propose the Most Frequency—Most Population (mFmP) heuristic that uses this association. However, the population is not the only way to estimate the prior probability. It can be inferred using other sources, such as number of links to Wikipedia (Overell and Ruger 2008) or, for example, the number of Instagram photos geotagged at that location. All these approaches can be seen as an adaptation of the Most Frequent Sense (MFS) heuristic used in general word sense disambiguation, where the most frequent sense is assigned to an ambiguous word (Calvo and Gelbukh 2014).

The first step in formulating a question is finding the best attribute on which the question will be based. The next step is selecting one or more appropriate feature values to be included in the question. If, for example, the best attribute is *Country* and the best attribute value is *United States*, the user can be asked whether they mean *San Antonio* in *United States*.

Questions can be of two types: yes–no questions (e.g., *Do you mean San Antonio in United States?*) and multiple-choice questions (e.g., *Do you mean San Antonio in USA, Costa Rica, or Chile?*). Multiple-choice questions are used only when the number of possible answers is low enough, so the user is not overwhelmed with too many options. This issue is known as Hick's Law in the UX domain; the more choices users face, the longer and harder it is for them to make a decision (Hick 1952).

The last step is question generation. Since there are a limited number of features, it is easy to define a small number of question templates that make the questions look natural. The templates are filled with the names and values of relevant features. If a question is based, e.g., on feature *Country*, the question template can look as follows: *Is <name> located in <continent>?* The placeholder *<name>* is then replaced by the name that is being disambiguated and *<continent>* with a continent name, for example, *Is Brno located in Europe?* There might be multiple templates for one question type to make the conversation less monotonic.

Based on the answer of a user, the candidates that have or do not have the specified property are filtered out. For example, if the user answers *no* when a question whether *San Antonio is in North America*, all places named *San Antonio* that are in North America will be eliminated.

The process of selecting a feature, its value, formulating a question, and eliminating candidates based on the answer is repeated until one candidate remains, or no further questions can be asked. Below we provide an example dialog between a user and the system showcasing the interactive

disambiguation process. In this example, when the user mentions *San Antonio*, the system lists possibilities and asks the user to narrow it down.

User: I am planning a trip to San Antonio. Could you recommend a book about its architecture?

System: There are multiple places named San Antonio. Are you referring to San Antonio in the USA? (*yes-no question*)

User: No, that is not the one I had in mind.

System: I see. Perhaps you meant San Antonio in Chile, Venezuela, or Bolivia? (*multiple-choice question*)

User: Yes, the one in Chile.

System: Understood, you need a book on the architecture of *San Antonio in Chile*. I would be happy to assist with that.

Question Formulation Strategies

We implemented and tested several approaches to both steps affecting the questions, i.e., selecting a suitable feature and its value. We used the baseline method (Equi-size partitioning), which is generally optimal when no additional information regarding the probability of the correct answer is available. The second baseline is *Population heuristics*. We also propose a novel approach called *Equi-population partitioning* that integrates aspects of existing methods.

Equi-Size Partitioning

The feature and its value are selected so splitting the set of potential candidates into approximately equal-sized subsets is enabled:

$$\operatorname{argmin}_{f \in F, v \in V_f} \left| |X_{f=v}| - \frac{|X|}{2} \right|$$

where X is the set of candidates, F is a set of features, $X_{f=v}$ is a set of candidates where the value of their feature f equals v , and V_f is a set of possible values of feature f .

When there are, e.g., 20 candidates, the question is selected so that answering it eliminates them to a number closest to 10 (one half of 20). For example, the question *Is <name> in Europe?* splits the candidates into two groups: places that are in Europe (9 candidates) and those that are not (11 candidates). Another question *Is <name> in the Czech Republic?* splits the candidates into two groups – places that are in the Czech Republic (2 candidates) and those that are not (18 candidates). The former question, which is based on feature *Continent* and its value *Europe*, is

better than the latter, which is based on feature *Country* and its value *Czech Republic*.

Population Heuristics

In this strategy, the population of the candidates is considered as the only available information for disambiguation. The candidates are presented to the user in the order given by their population in descending order. The number of questions that need to be asked is therefore equal to the order of the toponym within the list of candidates.

Equi-Population Partitioning

This new herein proposed approach leverages an Equi-partitioning strategy but enhances it through a measure of prior probability, which is in our case population. The feature and its value are selected so splitting the set of potential candidates into subsets with approximately equal populations is enabled:

$$\operatorname{argmin}_{f \in F, v \in V_f} \left| \operatorname{population}(X_{f=v}) - \frac{\operatorname{population}(X)}{2} \right|$$

where X is the set of candidates, F is a set of features, $X_{f=v}$ is a set of candidates where the value of their feature f equals v , V_f is a set of possible values of feature f , and *population* is a function returning the total population of a set of candidates.

When there are, e.g., 20 candidates with a total population of 1,000,000, the question is selected so that answering it eliminates the candidates to a set of a total population closest to 500,000 (one half of 1,000,000). For example, the question *Is <name> in Europe?* splits the candidates into two groups – places that are in Europe (12 places, total population of 700,000) and those that are not (8 places, total population of 300,000). Another question *Is <name> in the Czech Republic?* splits the candidates into two groups: places that are in the Czech Republic (10 places, total population of 100,000) and those that are not (10 places, total population of 900,000). The former question, which is based on feature *Continent* and its value *Europe*, is better than the latter, which is based on feature *Country* and its value *Czech Republic*.

EVALUATION

In this section, we evaluate our interactive toponym disambiguation methods. The core principle of using dialog clarification techniques for disambiguation (i.e., leveraging additional user input to resolve ambiguity) is established in various domains (Zamani et al. 2020; Braslavski et al. 2017; Shao et al. 2022; Alfieri, Wolter, and Hashemi 2022), including in database searches

(Qian et al. 2022) and can be generalized to the field of toponym disambiguation.

Our main focus is on the selection of features upon which questions are formulated. Therefore, we have not evaluated aspects such as geotagging, named entity recognition (NER), question generation, and user experience (UX), allowing us to concentrate our assessment solely on the methods for feature selection.

This evaluation is conducted fully automatically, utilizing a corpus containing real-world instances of toponym usage as a source of ground truth data. For this purpose, we developed an automated script, where the system attempts to disambiguate toponyms while the user responses (that are known in advance, since the correct toponym candidate is known) are simulated using toponym candidates from the corpus.

Evaluation Dataset

To perform the evaluation, we needed a list of place names users might want to disambiguate. Each name in this list must be ambiguous (i.e., there are multiple places with the same name) and the correct place is known (i.e., a reference to a specific place is needed). The frequency of occurrence of the place name candidates should also mirror the natural candidates' distribution to see how the proposed methods work in real scenarios where some candidates occur more often than others. Because no information about the frequency of toponym usage is available for a given domain, we relied on a naturally created text corpus containing annotated ambiguous place names. The content of the texts from the corpus is not important for the evaluation and the corpus is thus used only as a source of unambiguous toponyms.

For this reason, we have to exclude synthetically generated datasets, since the candidate occurrences do not follow real-world usage and focus on corpora composed of texts originating from human authors. We also excluded datasets with license restrictions. Since our system is built upon the Geonames gazetteer, we limited the evaluation to corpora that use Geonames identifiers to identify places. We combined the following corpora: (i) GeoCorpora (Wallgrün et al. 2018), predominantly composed of microblog texts, (ii) Local Global Corpus (Lieberman, Samet, and Sankaranarayanan 2010) containing news articles from geographically distributed newspapers, and (iii) GeoWebNews (Gritta, Pilehvar, and Collier 2020) with texts from globally distributed news sites. Our combined dataset contained 338 distinct place names with 4,199 candidate places in total. Each place name had at least three possible candidate references, requiring more than one question for disambiguation. The distribution of the number of candidates per place name is shown in [Figure 4](#).

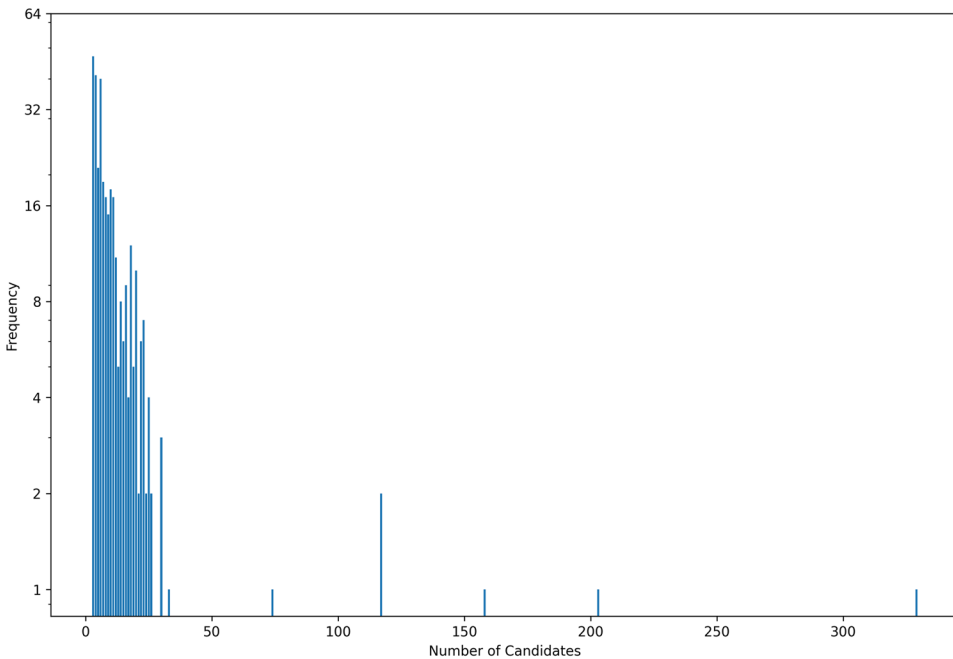


Figure 4. Frequency of toponyms for a given number of candidates.

Evaluation Methodology

For each ambiguous place name in the evaluation dataset, all entities with that name together with their properties were retrieved from Geonames. The correct candidate was known, constituting the ground truth. The obtained list of candidates was then eliminated based on the questions formulated using the properties of the entities until the correct candidate remained.

We compare three methods introduced in the Question Formulation Strategies section for choosing the feature on which the clarification question is asked: Equi-population partitioning, Equi-size partitioning, and a variant of the population heuristic (Rauch, Bukatin, and Baker 2003). We also evaluate the performance of these methods when using different types of questioning strategies. In one scenario, only yes-no questions are employed; in another scenario, we extend this approach by allowing selection from a list of candidates. We perform the evaluation for the number of options from which a user can choose (3, 4, and 5 options).

We measure the performance of each method by the *total number of questions* needed to disambiguate all toponyms. To eliminate the possible negative impact of place names with many candidates where we might expect slow disambiguation (requiring many questions), we relate the

Table 2. Performance of toponym disambiguation methods.

Yes-no questions (toponyms to disambiguate: 338, total number of candidates: 4199)			
Method	Total no. of questions asked	Success rate	Average no. of questions per candidate
Population heuristics	947	86.7	0.22
Equi-size partitioning	1,097	90.0	0.26
Equi-population partitioning	812	91.7	0.19
3 multiple-choice (toponyms to disambiguate: 291, total number of candidates: 4058)			
Method	Total no. of questions asked	Success rate	Average no. of questions per candidate
Population heuristics	875	84.5	0.22
Equi-size partitioning	936	92.4	0.23
Equi-population partitioning	727	92.8	0.18
4 multiple-choice (toponyms to disambiguate: 250, total number of candidates: 3894)			
Method	Total no. of questions asked	Success rate	Average no. of questions per candidate
Population heuristics	799	82.4	0.21
Equi-size partitioning	755	93.6	0.20
Equi-population partitioning	641	93.6	0.16
5 multiple-choice (toponyms to disambiguate: 229, total number of candidates: 3789)			
Method	Total no. of questions asked	Success rate	Average no. of questions per candidate
Population heuristics	742	82.4	0.20
Equi-size partitioning	665	93.4	0.18
Equi-population partitioning	573	94.3	0.15

Note: The bold values indicate the best results for the given metric.

number of questions asked to the number of candidates and present the *average number of questions per candidate*. We also present a supporting metric: *success rate*, which is defined as the fraction of samples disambiguated by asking five or fewer questions.

RESULTS

In this section, we present an evaluation of three methods across two distinct types of questioning strategies: yes-no, and options with the number of choices ranging from three to five. Table 2 summarizes the overall performance metrics for each method.

For the yes-no questioning approach, Equi-population partitioning performed the best across all measured metrics. Conversely, the Equi-size partitioning method exhibited the worst performance in terms of the total number of questions and the average number of questions per candidate.

When employing options-based questioning, i.e., allowing a choice from a list of candidates, Equi-population partitioning still performed the best across all metrics for all option sizes. The performance of Equi-size

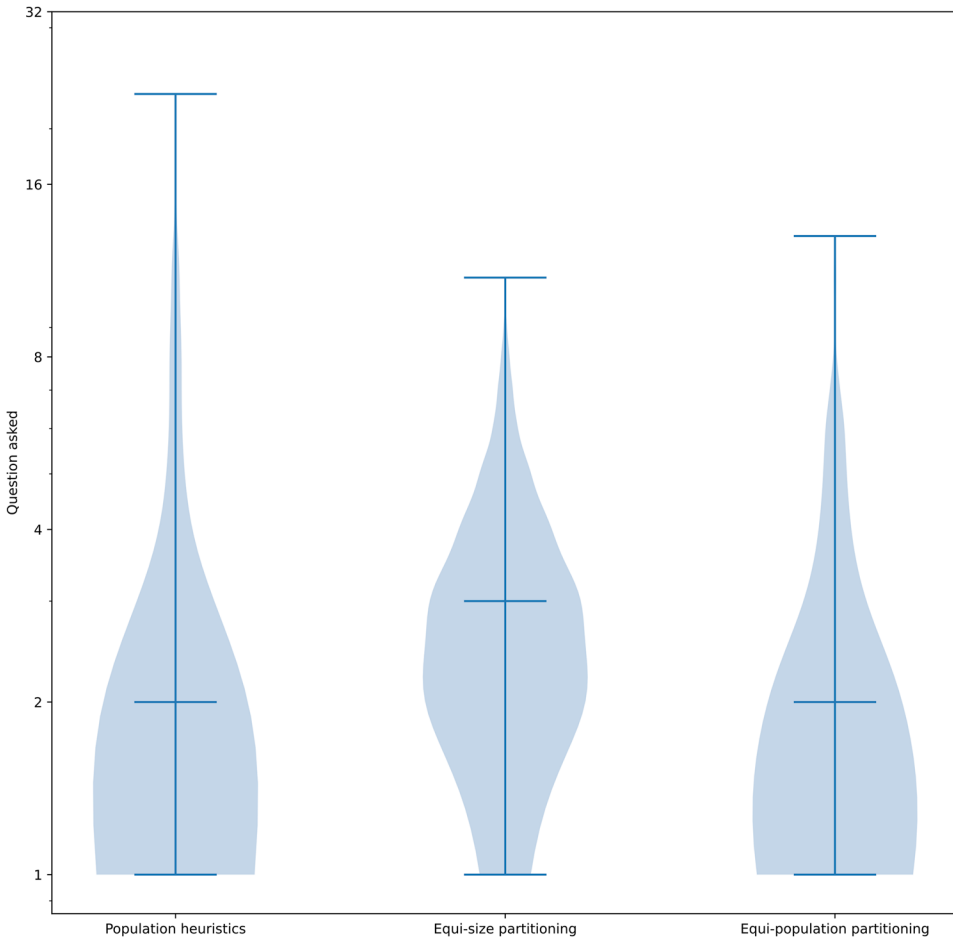


Figure 5. Distribution of the number of questions asked for each method—only yes-no questions.

partitioning and population heuristics varied with the number of answer options.

The distribution of the number of questions asked per toponym is depicted in [Figure 5](#). It illustrates that Equi-population partitioning performs similarly to the partitioning using population heuristics. The difference can be seen mostly when a larger number of questions needed to be asked (toponyms with many candidates). Here, Equi-population partitioning generally prevented asking too many questions.

[Figure 6](#) is a different representation of the same data and shows the average number of questions asked in relation to different numbers of candidates. It demonstrates how each method performs for a given number of candidates.

[Figure 7](#) shows the relation between the number of toponym candidates and the sum of the average number of questions that must be answered

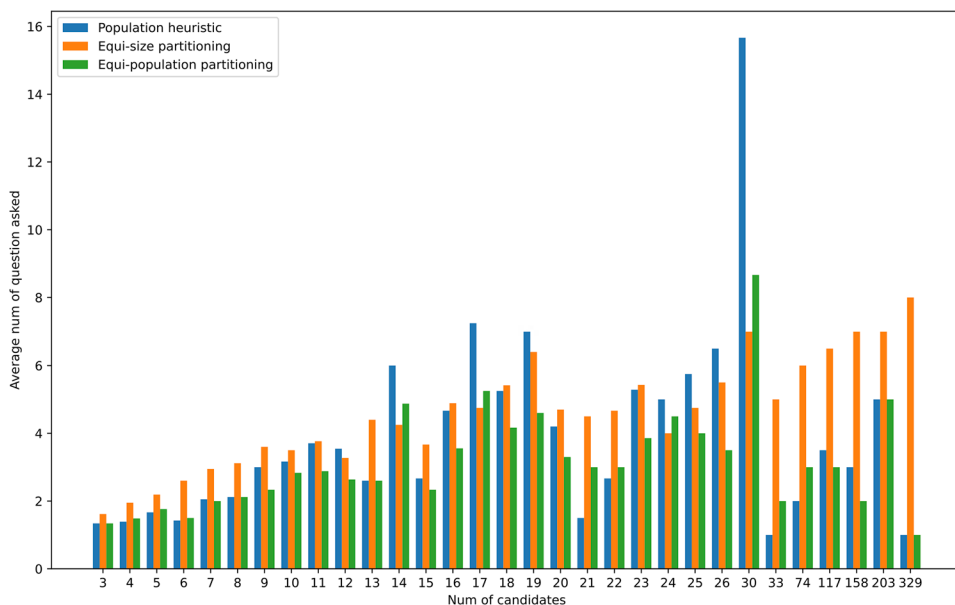


Figure 6. Cumulative number of questions asked—only yes-no questions.

to disambiguate the toponyms. This figure is intended to illustrate how each method performs with respect to the rising number of candidates. The lower the curve is located, the better the performance.

These findings indicate that the Equi-population strategy informed by the population of a candidate provided an efficient way to interactively disambiguate toponyms in a variety of different cases. We can conclude that additional information, in our case the population of candidates, generally improved the disambiguation process by reducing the number of questions asked – as Equi-population partitioning is an informed version of Equi-size partitioning, all three evaluation metrics reached their best values for the Equi-population partitioning method. Interestingly, without considering any external knowledge, the Equi-size method arrived at a correct answer using five or fewer questions more often than when using the population heuristics.

CONCLUSIONS

This contribution is poised to fill a gap in automated interactive place name disambiguation, offering a practical and efficient solution that aligns with the evolving demands of information retrieval in dialog systems.

Our research addresses a gap in toponym disambiguation methods suitable for the use in interactive systems by presenting a practical and efficient approach to toponym disambiguation. The proposed interactive approach, in which the user plays a crucial role in resolving ambiguity,

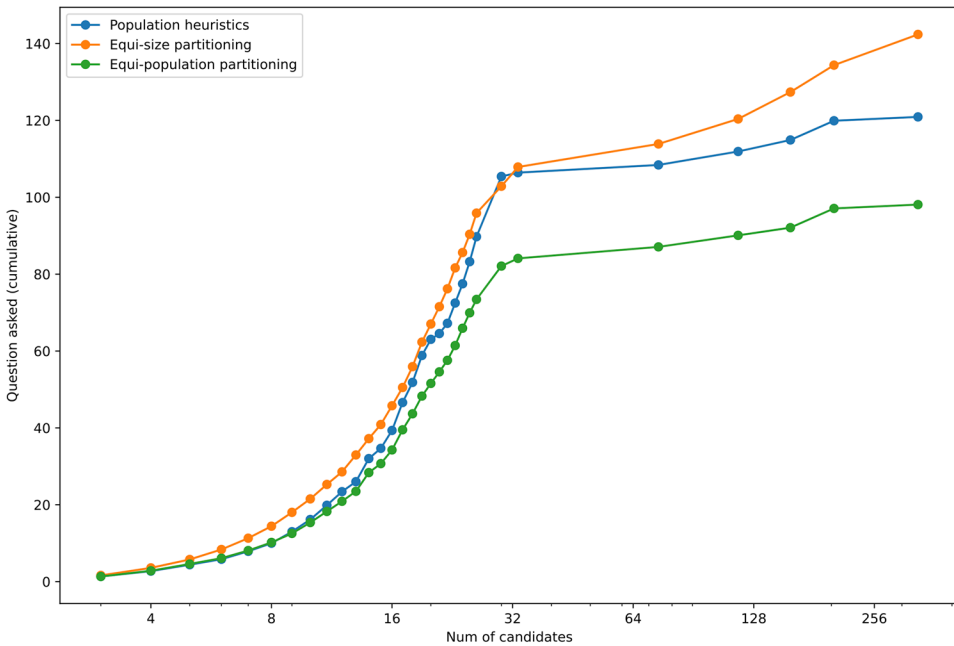


Figure 7. Average number of questions asked—only yes-no questions.

is well-suited for dialog systems, conversational interfaces, or in principle any interface that is interaction-limited since it exploits the interactivity and leverages additional information obtained directly from users to resolve ambiguity. It can find applications in chatbots or interactive systems that need to work with unambiguous spatial references. This may include applications in the domain of information retrieval, geospatial database search, geographic information systems, etc.

In particular, we have focused on methods for choosing the feature used to formulate a question that helps to narrow down ambiguous geographic references. We proposed a method called Equi-population partitioning, which combines the strengths of the population heuristic and the Equi-size partitioning method. This method is guaranteed to achieve the correct answer with the smallest number of questions when no source of prior probability is available. The proposed method allowed us to disambiguate place references appearing in real-world texts with fewer questions on average. All other supporting metrics (e.g., success rate or average number of questions per candidate) also showed improved performance.

Moving forward, several research questions remain to be explored. A comparative analysis of the effectiveness of using population versus other sources of prior probability is one future possibility. In addition, conducting a user study to evaluate the perception of our approach versus more traditional disambiguation approaches could provide valuable insight into the user experience.

DISCLOSURE STATEMENT

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

ORCID

Mikuláš Muroň  <http://orcid.org/0000-0003-1344-1772>

David Procházka  <http://orcid.org/0000-0002-2593-7495>

Roman Kern  <http://orcid.org/0000-0003-0202-6100>

REFERENCES

- Acheson, E., R. S. Volpi, and M. Purves. 2020. Machine learning for cross-gazetteer matching of natural features. *International Journal of Geographical Information Science* 34 (4):708–34. doi: [10.1080/13658816.2019.1599123](https://doi.org/10.1080/13658816.2019.1599123).
- Adamopoulou, E, and L. Moussiades. 2020. Chatbots: History, technology, and applications. *Machine Learning with Applications* 2:100006. doi: [10.1016/j.mlwa.2020.100006](https://doi.org/10.1016/j.mlwa.2020.100006).
- Alfieri, A., R. Wolter, and S. H. Hashemi. 2022. Intent disambiguation for task-oriented dialogue systems. *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 5079–5080. doi: [10.1145/3511808.3557516](https://doi.org/10.1145/3511808.3557516).
- Andogah, G. 2011. *Geographically constrained information retrieval: Geographically intelligent information retrieval*. Koln, DEU: LAP Lambert Academic Publishing.
- Boyarsky, K., D. Kanevsky, and Butorina, E. 2020. Automatic identification and classification of ambiguous Toponyms. *CEUR Workshop Proceedings* :285–98.
- Brando, C., and F. Frontini. 2017. Semantic historical gazetteers and related NLP and corpus linguistics applications. *Journal of Map & Geography Libraries* 13 (1):1–6. doi: [10.1080/15420353.2017.1307307](https://doi.org/10.1080/15420353.2017.1307307).
- Braslavski, P., D. Savenkov, E. Agichtein, and A. Dubatovka. 2017. What do you mean exactly? Analyzing clarification questions in CQA. *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*, 345–8.
- Buscaldi, D. 2010. Toponym disambiguation in information retrieval. Doctoral dissertation., Universidad Politecnica de Valencia.
- Buscaldi, D., and P. Rosso. 2008. Map-based vs. knowledge-based toponym disambiguation. *Proceedings of the 5th Workshop on Geographic. Information Retrieval*, 19–22. doi: [10.1145/1460007.1460011](https://doi.org/10.1145/1460007.1460011).
- Butler, J. O., J. E. Donaldson, I. N. Taylor, and C. E. Gregory. 2017. Alts, Abbreviations, and AKAs: Historical onomastic variation and automated named entity recognition. *Journal of Map & Geography Libraries* 13 (1):58–81. doi: [10.1080/15420353.2017.1307304](https://doi.org/10.1080/15420353.2017.1307304).
- Cai, G., H. Wang, A. M. MacEachren, and S. Fuhrmann. 2005. Natural conversational interfaces to geospatial databases. *Transactions in GIS* 9 (2):199–221., and doi: [10.1111/j.1467-9671.2005.00213.x](https://doi.org/10.1111/j.1467-9671.2005.00213.x).
- Calvo, H., and A. Gelbukh. 2014. Finding the most frequent sense of a word by the length of its definition. *Human-Inspired Computing and Its Applications* :1–8.
- CBC News. 2010. Italian tourists end up in wrong Sydney. Available at <https://www.cbc.ca/news/canada/nova-scotia/italian-tourists-end-up-in-wrong-sydney-1.975605>. (accessed March 29, 2024).
- Chakrabarti, B., and P. Mahapatra. 1989. *Library and information science: An introduction*. World Press Private Limited.

- Coden, A., D. Gruhl, N. Lewis, and P. N. Mendes. 2015. *Did you mean A or B? Supporting clarification dialog for entity Disambiguation. Proceedings of the 1st International Workshop on Summarizing and Presenting Entities and Ontologies and the 3rd International Workshop on Human Semantic Web Interfaces.*
- Dagan, Y., Y. Filmus, A. Gabizon, and S. Moran. 2017. Twenty (simple) questions. 49th ACM Symposium on Theory of Computing. doi: [10.1145/3055399.3055422](https://doi.org/10.1145/3055399.3055422).
- DeeAnn, A. 2012. Chatbots in the library: Is it time? *Library Hi Tech* 30 (1):95–107. doi: [10.1108/07378831211213238](https://doi.org/10.1108/07378831211213238).
- Ehrenpreis, M., and J. DeLooper. 2022. Implementing a Chatbot on a Library Website. *Journal of Web Librarianship* 16 (2):120–42. doi: [10.1080/19322909.2022.2060893](https://doi.org/10.1080/19322909.2022.2060893).
- El Midaoui, O., B. El Ghali, A. El Qadi, and M. D. Rahmani. 2018. Geographical query reformulation using a geographical taxonomy and WordNet. *Procedia Computer Science* 127:489–98. doi: [10.1016/j.procs.2018.01.147](https://doi.org/10.1016/j.procs.2018.01.147).
- ESRI. 2018. Sonar. GitHub repository, Available at: <https://github.com/Esri/sonar>. (accessed March 29, 2024)
- Evans, W. and D. Baker, eds. 2011. *Libraries and society: Role, responsibility and future in an age of change*. Amsterdam: Elsevier.
- Fu, J., S. Mouakket, and Y. Sun. 2023. The role of chatbots' human-like characteristics in online shopping. *Electronic Commerce Research and Applications* 61:101304. doi: [10.1016/j.elerap.2023.101304](https://doi.org/10.1016/j.elerap.2023.101304).
- Garrel, J., and J. Mayer. 2023. Artificial intelligence in studies—use of ChatGPT and AI-based tools among students in Germany. *Humanities and Social Sciences Communications* 10 (1):1–9. doi: [10.1057/s41599-023-02304-7](https://doi.org/10.1057/s41599-023-02304-7).
- Ghufran, M., G. Quercini, and N. Bennacer. 2015. Toponym disambiguation in online social network profiles. GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems: a6. doi: [10.1145/2820783.2820812](https://doi.org/10.1145/2820783.2820812).
- Gritta, M., M. T. Pilehvar, and N. Collier. 2020. A pragmatic guide to geoparsing evaluation. *Language Resources and Evaluation* 54 (3):683–712., and doi: [10.1007/s10579-019-09475-3](https://doi.org/10.1007/s10579-019-09475-3).
- Habib, M. B., and M. Van Keulen. 2016. TwitterNEED: A hybrid approach for named entity extraction and disambiguation for tweet. *Natural Language Engineering* 22 (3):423–56. doi: [10.1017/S1351324915000194](https://doi.org/10.1017/S1351324915000194).
- Hadi Mogavi, R., C. Deng, J. Juho Kim, P. Zhou, Y. D. Kwon, A. Hosny Saleh Metwally, A. Tlili, S. Bassanelli, A. Bucchiarone, S. Gujar, et al. 2024. ChatGPT in education: A blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. *Computers in Human Behavior: Artificial Humans* 2 (1):100027. doi: [10.1016/j.chbah.2023.100027](https://doi.org/10.1016/j.chbah.2023.100027).
- Hick, W. E. 1952. On the rate of gain of information. *Quarterly Journal of Experimental Psychology* 4 (1):11–26. doi: [10.1080/17470215208416600](https://doi.org/10.1080/17470215208416600).
- Hu, H., X. Wu, B. Luo, C. Tao, C. Xu, W. Wu, and Z. Chen. 2018. Playing 20 question game with policy-based reinforcement learning. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing
- Kafando, R., R. Decoupes, M. Roche, and M. Teisseire. 2023. SNEToolkit: Spatial named entities disambiguation toolkit. *SoftwareX* 23:101480. doi: [10.1016/j.softx.2023.101480](https://doi.org/10.1016/j.softx.2023.101480).
- Karimzadeh, M., W. Huang, S. Banerjee, J. O. Wallgrün, F. Hardisty, S. Pezanowski, P. Mitra, and A. M. MacEachren. 2013. GeoTxt: A web API to leverage place references in text. *Proceedings of the 7th Workshop on Geographic Information Retrieval*, 72–3.
- Kejriwal, M., and P. Szekely. 2017. Neural embeddings for populated geonames locations. International Semantic Web Conference, 139–146.

- Khan, W., A. Daud, K. Khan, S. Muhammad, and R. Haq. 2023. Exploring the frontiers of deep learning and natural language processing: A comprehensive overview of key challenges and emerging trends. *Natural Language Processing Journal* 4:100026. doi: [10.1016/j.nlp.2023.100026](https://doi.org/10.1016/j.nlp.2023.100026).
- Lafia, S., J. Xiao, T. Herve, and W. Kuhn. 2019. Talk of the town: Discovering open public data via voice assistants. 14th International Conference on Spatial Information Theory.
- Leidner, J. L. 2007. Toponym resolution in text: Annotation, evaluation and applications of spatial grounding. *ACM SIGIR Forum* 41 (2):124–6. doi: [10.1145/1328964.1328989](https://doi.org/10.1145/1328964.1328989).
- Lieberman, M. D., and H. Samet. 2012. Adaptive context features for toponym resolution in streaming news. *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, 731–740. doi: [10.1145/2348283.2348381](https://doi.org/10.1145/2348283.2348381).
- Lieberman, M. D., H. Samet, and J. Sankaranarayanan. 2010. Geotagging with local lexicons to build indexes for textually-specified spatial data. *2010 IEEE 26th International Conference on Data Engineering*, 201–12.
- McNeal, M. L., and D. Newyear. 2013. Introducing Chatbots in Libraries. *Library Technology Reports* 49 (8):5–10.
- Middleton, S. E., S. Kordopatis-Zilos, Y. Papadopoulos and, and Kompatsiaris, G. 2018. Location extraction from social media: Geoparsing, location disambiguation, and geotagging. *ACM Transactions on Information Systems* 36 (4):1–27. doi: [10.1145/3202662](https://doi.org/10.1145/3202662).
- Nawaz, N., and M. A. Saldeen. 2020. Artificial Intelligence Chatbots for Library Reference Services. *Journal of Management Information and Decision Sciences* 23:442–9.
- Overell, S., and S. Rürger. 2008. Using co-occurrence models for placename disambiguation. *International Journal of Geographical Information Science* 22 (3):265–87. doi: [10.1080/13658810701626236](https://doi.org/10.1080/13658810701626236).
- Panda, S., and R. Chakravarty. 2022. Adapting intelligent information services in libraries: A case of smart AI Chatbots. *Library Hi Tech News* 39 (1):12–5. 2022. doi: [10.1108/LHTN-11-2021-0081](https://doi.org/10.1108/LHTN-11-2021-0081).
- Purver, M., J. Ginzburg, and P. Healey. 2003. On the means for clarification in dialogue. *Current and New Directions in Discourse & Dialogue*. 235–55.
- Qi, T., S. Ge, C. Wu, Y. Chen, and Y. Huang. 2019. Toponym Detection and Disambiguation on Scientific Papers. *Proceedings of the 13th International Workshop on Semantic Evaluation* 1302–1307. doi: [10.18653/v1/S19-2229](https://doi.org/10.18653/v1/S19-2229).
- Qian, K., S. Kottur, A. Beirami, S. Shayandeh, P. Crook, A. Geramifard, Z. Yu, and C. Sankar. 2022. Database Search Results Disambiguation for Task-Oriented Dialog Systems. *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 1158–1173. doi: [10.18653/v1/2022.naacl-main.85](https://doi.org/10.18653/v1/2022.naacl-main.85).
- Radio, E., S. Gregonis, P. Werling, K. White, and K. Crosbi, 2021. Evaluation of geographic vocabularies and their usage. *Journal of Map & Geography Libraries* 17 (1):1–23. doi: [10.1080/15420353.2022.2068108](https://doi.org/10.1080/15420353.2022.2068108).
- Rauch, E., M. Bukatin, and K. Baker. 2003. A confidence-based framework for disambiguating geographic terms. *Proceedings of the HLT-NAACL 2003 Workshop on Analysis of Geographic References*, 50–4.
- Reinsfelder, T. L., and K. O'Hara-Krebs. 2023. Implementing a rules-based chatbot for reference service at a large university library. *Journal of Web Librarianship* 17 (4):95–109. doi: [10.1080/19322909.2023.2268832](https://doi.org/10.1080/19322909.2023.2268832).
- Rodriguez, S., and C. Mune. 2022. Uncoding library Chatbots: deploying a new virtual reference tool at the San Jose state university library. *Reference Services Review* 50 (3/4):392–405. doi: [10.1108/RSR-05-2022-0020](https://doi.org/10.1108/RSR-05-2022-0020).

- Sanji, M., G. Behzadi, and H. Gomroki. 2022. Chatbot: An intelligent tool for libraries. *Library Hi Tech News* 39 (3):17–20. doi: [10.1108/LHTN-01-2021-0002](https://doi.org/10.1108/LHTN-01-2021-0002).
- Santos, R., P. Murrieta-Flores, B. Calado, and P. Martins. 2018. Toponym matching through deep neural networks. *International Journal of Geographical Information Science* 32 (2):324–48. doi: [10.1080/13658816.2017.1390119](https://doi.org/10.1080/13658816.2017.1390119).
- Scaramozzino, J., R. White, J. Essic, L. A. Fullington, H. Mistry, A. Henley, and M. Olivares. 2014. Map room to data and GIS services: Five university libraries evolving to meet campus needs and changing technologies. *Journal of Map & Geography Libraries* 10 (1):6–47., and doi: [10.1080/15420353.2014.893943](https://doi.org/10.1080/15420353.2014.893943).
- Shannon, C. E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27 (3):379–423. doi: [10.1002/j.1538-7305.1948.tb01338.x](https://doi.org/10.1002/j.1538-7305.1948.tb01338.x).
- Shao, T., C. Fei, C. Wanyu, and C. Honghui. 2022. Self-supervised clarification question generation for ambiguous multi-turn conversation. *Information Sciences* 587:626–41. doi: [10.1016/j.ins.2021.12.040](https://doi.org/10.1016/j.ins.2021.12.040).
- Soo, V. W., and H. L. Cheng. 2002. Conducting the disambiguation dialogues between software agent sellers and human buyers. *PPacific Rim International Workshop on Multi-Agents*, 123–37. Berlin, Heidelberg, Springer.
- Suhaili, S. M., M. N. Salim, and N. Jambli. 2021. Service chatbots: A systematic review. *Expert Systems with Applications* 184:115461. doi: [10.1016/j.eswa.2021.115461](https://doi.org/10.1016/j.eswa.2021.115461).
- Suresh, S. R. 2017. A bayesian strategy to the 20 question game with applications to recommender systems., Doctoral dissertation., Duke University.
- Torvik, V. I. 2015. MapAffil: A bibliographic tool for mapping author affiliation strings to cities and their geocodes worldwide. *D-Lib Magazine* 21 (11/12):11–2. doi: [10.1045/november2015-torvik](https://doi.org/10.1045/november2015-torvik).
- Vincze, J. 2017. Virtual reference librarians (Chatbots). *Library Hi Tech News* 34 (4):5–8. doi: [10.1108/LHTN-03-2017-0016](https://doi.org/10.1108/LHTN-03-2017-0016).
- Wallgrün, J. O., M. Karimzadeh, A. M. MacEachren, and S. Pezanowski. 2018. GeoCorpora: Building a corpus to test and train microblog geoparsers. *International Journal of Geographical Information Science* 32 (1):1–29. doi: [10.1080/13658816.2017.1368523](https://doi.org/10.1080/13658816.2017.1368523).
- Zamani, H., S. Dumais, N. Craswell, P. Bennett, and G. Lueck. 2020. Generating clarifying questions for information retrieval. *Proceedings of the Web Conference 2020*, 418–28.
- Zhou, M., S. Duan, H. Y. Liu, and N. Shum. 2020. Progress in neural NLP: Modeling, learning, and reasoning. *Engineering* 6 (3):275–90. doi: [10.1016/j.eng.2019.12.014](https://doi.org/10.1016/j.eng.2019.12.014).