2015

# A Natural Language Search Interface for Accommodation Queries

Erin Chavanne
*Claremont McKenna College*

Claremont McKenna College

# A Natural Language Search Interface for Accommodation Queries

submitted to
Professor Arthur Lee
and
Professor America Chambers
and
Dean Nicholas Warner

by
Erin Elizabeth Chavanne

for
Senior Thesis
Spring 2015
April 27, 2015

## Abstract

Services that once required human interaction are now completed with the click of a few buttons. In general, this allows for a more streamlined process for activities such as sending messages (email or text messages), filing taxes, or even shopping for groceries. In terms of searching for hotels and travel accommodations however, this process has not proven to be the most effective as the speed and efficiency is hindered by the interface through which this information is available. Choosing a travel specific site, filling in the required fields, combing through results for the desired specifications, and then possibly repeating the process elsewhere, does not provide the ability for the user to express the entirety of their preferences for the accommodation and is therefore not an effective method for searching.

Natural language search provides a more accessible and intuitive interface for accommodation searching. Instead of specifying fields that may not encompass the the entirety of the desired search, the user is able to express all of the aspects in a single, natural language, search.

In this project, we propose a natural language search interface for accommodations such as hotels, hostels, or apartments. Data acquired through Amazon Mechanical Turk is used to create a system for extracting various accommodation fields. Zilyo and Expedia APIs are then queried for real-time accommodation listings. These results are then adjusted based on the specifics of the search that were not included in the original query. A natural language search of this kind is not only more accessible on the consumer end, but provides data that pertains directly to the the entirety of the intended search.

## Acknowledgments

I would like to thank my advisors, Professors America Chambers and Arthur Lee, for their guidance and encouragement. Additionally, I would like to thank my friends and family for their love and support.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Most travel search engines allow users to specify values for predetermined fields. This is perfectly suited for searches that do not have a wide range of variety. When searching for a service or product with multiple options however, we believe natural language searches are more fitting. Natural language searches allow the user to type out their entire search and are not limited by specific fields and options. We believe that a user will receive more precise search results if they formulate a natural language query that encompasses all aspects of their search, which is then parsed, searched against various databases, evaluated for it's relevance to the search, and returned to the user.

We focus on accommodation searches because natural language options are not currently widespread. Most travel search engines rely on filling in predetermined fields. These tend to be fields such as location, dates, and the number of people. For searches where a user's desired accommodation is more specific than these fields allow, a natural language search is perfectly suited. A natural language search will return more precise results in this field as it will allow more aspects of the query to be included in the search, allowing the engine to do more of the filtering instead of passing that responsibility on to the user.

This project proposes the creation of a natural language interface for accommodation searches with the goal of allowing for searches that encompass the entirety of the user's request. In order to accomplish this, we built a natural language interface. Because data for this type of search is not readily available, Amazon Mechanical Turk (AMT) is used to create a dataset comprised of human-created natural language searches specifically for accommodations. Models based off of this data are used to extract specific parameters of the search, such as the type of accommodation, location, and date. These fields are used to query APIs for Zilyo and Expedia which contain two distinct sets of real-time accommodation listings. Expedia allows access to hotel information, while Zilyo contains alternate types such as hostels and apartments. In order to encompass other aspects of the search, these results are filtered and the rankings are adjusted depending on their similarity to the search. In order to evaluate the results, the extraction models are evaluated against a held-out set of data, and we conducted a user study to determine how it performs in comparison to Expedia's web interface. In section 2, we present relevant work in the fields of natural language and travel-specific searches. Section 3 gives a detailed description of the system, including how data was collected, specifics for the field extraction models, use of the APIs and their results, and the user interface. Section 4 outlines the results acquired through the user test and from the extractions. Finally, the last two sections address the conclusions drawn from this project and areas for future work.

# Chapter 2

# Related Work

The field of natural language travel searches is still in it's early stages. While some search engines such as Google and Yahoo (which take natural language queries) have access to travel related information and will return flight information, searches for accommodations are still more successful through a specific travel site. The majority of these travel specific sites do not accept natural language queries. Therefore, research on the topic of a natural language search engine for travel related searches takes two forms: the development of natural language search engines and travel specific searches. The field of natural language search engines is evolving from the desire to create search engines that return more precise and accurate results to users. In terms of travel related searches, some research has been done into how people search for accommodations and how searches are constructed.

## 2.1 Search Engines

With the growing necessity for search engines, research is being conducted on how humans interact with search engines and how these trends are changing over time. [JS06] analyze the difference in the number of "one query sessions" over the time span from 1997 to 2002. They show that the number of users conducting a single search before moving on (as opposed to conducting several searches in the same session) has remained relatively stable, indicating that queries are not becoming more complex. When looking at the percent of users who only view one page of results (over the same time period), there is an indication that web searches are becoming less complex as this percentage is increasing. It could be argued though that this indicates an increase in user's skills and the ability to create more complex searches that return fewer, more accurate results. Their analysis of query logs show that 12 - 24% (depending on the search engine) of all searches between the late 1990s and early 2000s were for commerce, travel, employment, or economy. This indicates that a significant portion of all searches are on travel, showing the necessity for search engines that are well-equipped for the field.

There have been several patents issued in the last 15 years regarding the technology behind natural language meta-searches. A meta-search is when a user inputs a single request which is in turn searched against more than one database, and the results are combined and returned to the user. [Red98] describes a procedure for this type of search that includes parsing a natural language query into queries for a specific selection of other search engines, filtering the results based on the query, ranking the results, and returning to the user. Filtering is used to remove irrelevant results. Ranking is used to process the results and select the most relevant based on the query terms and importance. In this project we search multiple databases in order to acquire data for a variety of searches, such as hotels versus apartments or hostels which may be accessible through different sources.

## 2.2 Natural Language Searches

There also exists research comparing the difference in results of natural language queries to those of boolean searches and the potential for higher precision using natural language. A boolean search is where a user inputs parts of the request into predetermined fields.

[Tur94] uses legal text documents to show that natural language queries are at least as efficient as boolean queries. Written in the early 1990s, it claims that commercial services are slow to implement natural language queries even though they are no less efficient. This research also discusses the process of comparing the boolean and natural language results and the challenges of comparing the two sets of search results. The results of the experiment show that natural language queries perform better for all except the most basic searches. Additionally, a natural language query will return more relevant results and present them in a more accurate order, such that the "best" result for the specific search is returned first.

[BHCDV99] address a method for post-processing search results using natural language methods. They propose to use natural language processing to produce a set of logical terms for both the search and each result. The "logical form" for the original search and for each result would be compared and scored based on their similarity. These new scores would be used to rank the results to be returned to the user. This method can be used to capture some of the semantic meaning in the text that is missed when extracting fields to create a search.

Strzalkowski et al. researched the concept of using natural language processing in the pre-processing stage. This research argues that the current methods for data retrieval are not robust enough and that natural language processing can be used to create more accurate searches. Instead of matching terms from the search to terms in the data, this proposes to use natural language processing to create a "representation" of the data which is then searched. This "representation" is a complex combination of extracted phrases that have been deemed semantically significant. The search is also parsed for semantically significant terms and phrases, and then the data is searched. This sort of natural language processing incorporation also allows more components of the search to be included, creating more semantic-focused searches. This provides motivation to create natural language search engine accessibility for queries of all types.

## 2.3 Travel Searches

The other aspect of this project is the focus on travel specific searches. [XP11] looks at patterns in travel queries, particularly their construction. This research suggests that there are three types of goals for a query: either navigational, informational, or transactional. The first being the desire to find a specific web page, the second looking for information on a specific subject, and third being an attempt to complete a transaction (i.e. purchase specific airline tickets). Research in this topic has lead to conclusions on how a typical user searches for travel information. Particularly, that the majority of users will generate general queries instead of specific ones, and that queries focus on quantifiable attributes (price, location, etc.). Additionally, users queries are highly reflective of the knowledge they have of a particular location. Other results include information on the use of keywords/terms (not including destination terms, only 372 other terms were used) and the length of search sessions (multiple vs. single queries).

This project builds on the foundational research regarding the success of natural language searches and applies it to the research done on travel specific searches. The results of research into travel related queries appear to be based on searches that are done using the common form of travel related searching, resulting in the focus on quantifiable attributes that can be captured by individual fields, and the low number of descriptive terms used in searches. Ideally, the results of this project will cause these statistics to change, allowing for fewer but more descriptive searches.

# Chapter 3

# System Description

This section describes the system created for natural language travel search queries. Natural language sentences were collected through the use of Amazon Mechanical Turk. The natural language text was then parsed for fields such as accommodation type, location, date, number of people, price range, and rating ranges, and were extracted from the search. These fields were used to search Expedia and Zilyo using their APIs. Finally, the results and their ranking were adjusted based on their similarity to the original search to take into account the aspects that could not be searched initially. See Figure 3.1 for a diagram of the system.

## 3.1 Data

There exists no released dataset available specifically for natural language travel search queries. Therefore, Amazon Mechanical Turk (AMT) was used to generate data. AMT allows tasks to be posted to a service which then releases these tasks to a network of workers. Data was collected in 4 separate batches, each with 300 unique tasks. Each task provided the worker with several pieces of data corresponding to fields that would normally be included in a travel request. These pieces of data were changed for each task. These included fields for the location, dates, number of people, and amenities. The location was

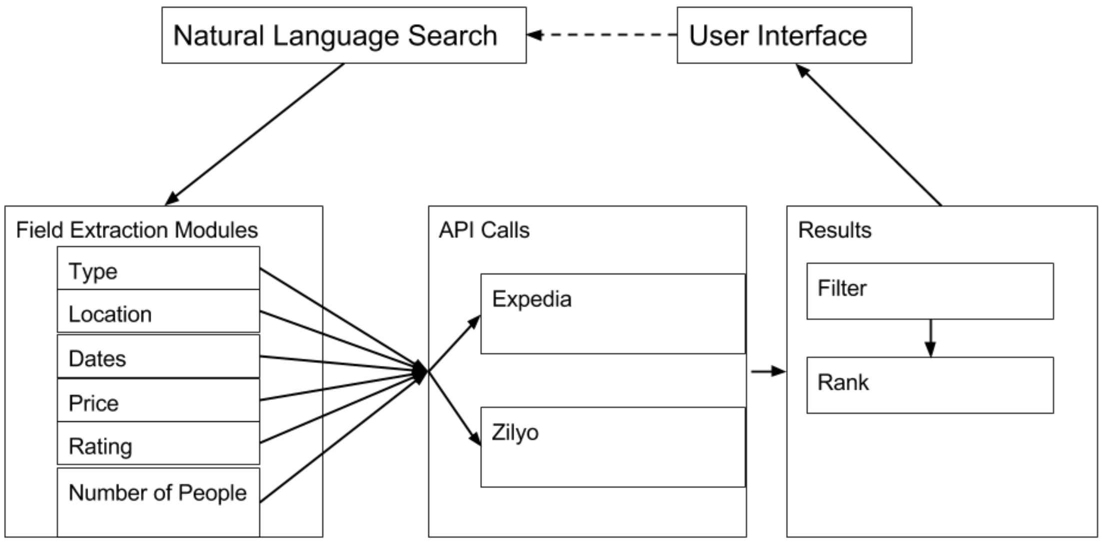


Figure 3.1: System Description

pulled from a list of all cities and counties in the United States. The type of accommodation was randomly generated from a list of possible accommodation types (hotel, apartment, and hostel). From the type of accommodation, the dates (and therefore length of stay) and price were generated for what would be reasonable given the type of accommodation. This means that a stay in an apartment might be longer than a hotel. The number of people and amenities were also randomly generated for each individual task.

The instructions then asked the user to write a few sentences using these pieces of data, as if they were speaking to a travel agent. The purpose of this was to encourage workers to form a request most like what they would ask another person, and less in the format that they may be used to when searching for travel information. Figure 3.2 gives an example of what a worker would see as their task. Additionally, it was requested that the worker be a native English speaker, and complete no more than 5 of these tasks. Other restrictions

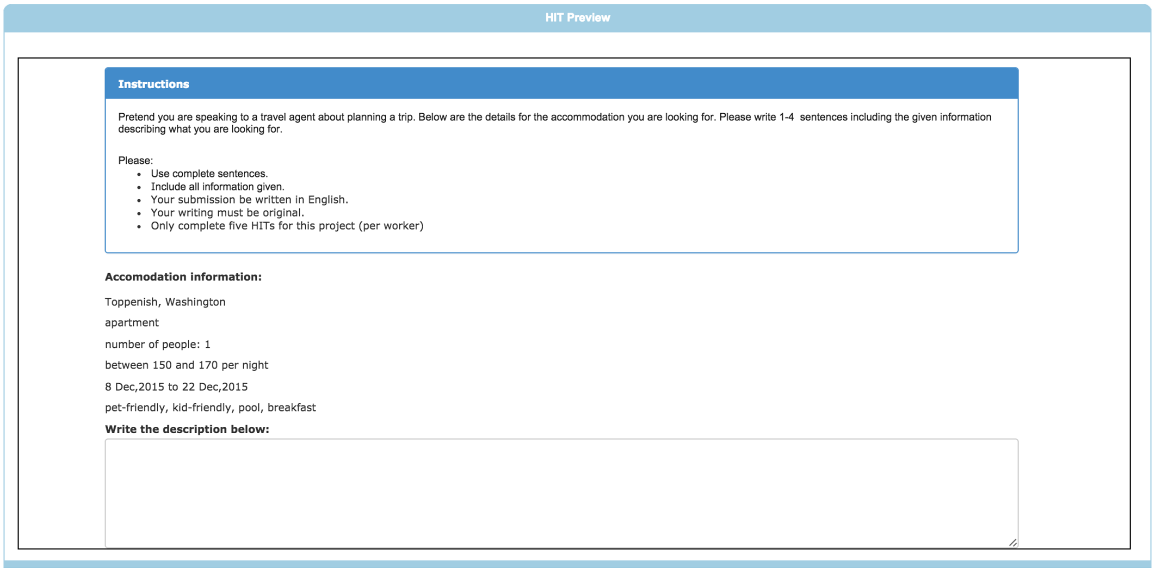


Figure 3.2: AMT Task

placed on the worker were that they had at most 1 hour to complete the task (although most workers completed it in less than a couple of minutes), that they have over a 90% approval rating for their previously completed tasks, and that they had at least 500 prior approved "HITS". These requirements were aimed to restrict the tasks to workers who had fairly high qualifications, but not so high as to severely limit the pool of workers. The structure of AMT's requests made it impossible to enforce that the workers were native English speaking and only completed 5 tasks without manually rejecting the worker's submission after it was completed. There was also a problem with workers completing more than the requested number of tasks or not following the instructions. However, since the goal of this project was to be able to handle a wide variety of natural language sentences (such as different wordings, limited information, or even recognizing off-topic searches), this resulted in acceptable data.

The tasks were completed in 4 batches, the first three of which were used for training and the last was used for testing. Of the 900 tasks released in the first 3 batches, 878 were returned. To account for a single worker completing more than the requested number of tasks, at most 15 tasks from a single worker were included. Any more than 15 from a unique worker ID were discarded. The purpose of this was to retain as many unique and potentially different natural language sentences without diluting the data if one worker formulated each of their queries similarly. This accounted for 383 rejections out of the original 878 submissions. Additionally, 20 more were manually rejected or filtered out for being too long or too short. This resulted in 495 natural language queries that were used to train the model. The same process was applied to clean the data from the 4th batch. This resulted in 110 out of 300 natural language search queries held out for testing.

## 3.2 Field Extractions

In order to search for valid hotel and apartment listings, it was necessary to next extract those fields out of the natural language query. For one of the fields, there existed module available for use. For the remaining fields, the data was analyzed for patterns in how humans formulate requests for each specific field.

### 3.2.1 Accommodation Type

Knowing the type of accommodation requested is necessary for knowing which database or API to search for valid results. This field required simply determining whether certain keywords were used. To indicate that a hotel was being requested, the term "hotel" was searched for. Similarly for an apartment or hostel, those keywords were searched for within the query. Out of the 495 valid queries, 268 matched with "hotel", 201 returned as a non-hotel accommodation, and for 26 no accommodation type was found. In these 26 instances, the user either didn't indicate a preference, or misspelled the accommodation type. In these cases, accommodations of all types were searched for.

### 3.2.2 Location

Extracting the location specified in the query was done using the Stanford Named Entity Recognizer (NER) developed by the Stanford Natural Language Processing Group. The NER labels natural language text with a variety of tags, including “person”, “organization”, and “location.” The “location” feature was used to extract location from the queries. If no location was found, a default location was assigned for the purpose of conducting the search. The default location was the user’s physical location. The location was both specified and correctly detected in 92% of the training data.

### 3.2.3 Date

| Pattern | Example | Percent Matched |
|---|---|---|
| mm/dd/yy | I’m looking for an apartment for 4 in Athens County, Ohio for 2/8/15 until 2/21/15. I can’t pay more than 264/night but would like at least a 1 star room. It needs to be pet-friendly and have a pool. | 10/495 = 2.02% |
| month day year | I want a hotel room for 3 people in Greenwood County, South Carolina. Check in on March 23, 2015 and check out March 30,2015. The cost should be between 109 and 197 per night. The hotel should be kid-friendly, have a TV and wifi | 215/495 = 43.43% |
| day month year | We are searching for a hotel room in Valley Falls, South Carolina for 4 people. Our budget is less than $205 per night. Our stay is scheduled for 13 Jun, 2015 to 15 Jun, 2015. We would really like a place that offers a pool, a gym, parking and TV | 182/495 = 36.77% |
| month day-day year | I am looking for a hotel with a pool, breakfast, kitchen, and laundry facilities in Erskine, Minnesota. It needs to be at least 3 stars, between $184 and $208 per night for 4 people. I need to book the room from September 18-21, 2015 | 8/495 = 1.61 % |
| () (days or nights or weeks) starting (date) | Please arrange me an apartment for nine days from 21st December 2015 at Suttons Bay, Michigan with a 4 star or higher convenience to accommodate 2 persons. Following conveniences are must at the apartment: Gym, smoke-free rooms and a TV | 45/495 = 9.10% |
| () to the () of month | I’m looking for an apartment in Nashville, Tennessee for three people. I need it to be less than $220 per night from the dates of August 8th to the 20th 2015. It needs to also include a gym. | 3/495 = 0.61% |

Table 3.1: Patterns for the Date

Date extraction was done by identifying several patterns used to specify dates in the training data. The majority included sentences where the date was specified in a format such as “July 6th” or, “the 6th of July” and optionally including the year. These two patterns equated to 397 matches of the 496 training sentences. Additionally, patterns were found for specifying dates in the traditional “mm/dd/yy” or “mm/dd/yyyy” format, again allowing for the year to be optionally specified, defaulting to the current year. A third general form that was included was sentences where the initial date was specified as well as a number of nights, days, or weeks. For example, in the search “we will be there for 4 nights starting April 15th, 2015,” the date April 15th, 2015 would be recognized. The match to “4 nights” would add four days on to the original date to produce a valid date range. Similarly, if there was a match on “weeks,” the corresponding number of days would be added to the start date. If dates were not found, the default date range of today’s date and tomorrow’s date (a 1 day duration starting on the current date) was returned, allowing for database searches to be carried out. See Table 3.1 for patterns used and example sentences.

### 3.2.4 Number of People

| Pattern | Example | Percent Matched |
|---|---|---|
| () (people or person or persons or guest or guests) | Our trip destination is Sacaton, Arizona arriving on Sept 21, 2015 and departing on Oct 2, 2015. We need an apartment for 4 people in the price range of 285 to 354 per night. The accommodations must be pet-friendly, kid-friendly with a kitchen and access to a gym | 363/495 = 73.33%S |
| () of us | There are two of us looking for a hotel in Breezewood, Pennsylvania for the dates 20 Aug 2015 to 24 Aug 2015. We would prefer a hotel that is pet-friendly, has a pool, and at least a 3 star rating | 15/495 = 3.03% |
| () (members or family members) | I would like to go for a trip to Witten in South Dakota. I have planned to stay in a hotel for 5 nights. i.e. from 9 Apr, 2015 to 13 Apr, 2015. Please book a room for 3 family members. The facilities I required is gym, healthy breakfast, smoke-free rooms and TV | 16/495 = 3.23% |
| (myself or alone of just me) | I’m looking for an apartment in Fort Sill, Oklahoma for myself. I will be staying from 1 Oct,2015 to 16 Oct,2015 and will pay less than 214 per night. I need pet-friendly apartment, pool, gym, a-c | 9/495 = 1.81% |
| (family of or group of or party of) () | I’m looking for a hotel located in Ellendale, North Dakota.I’ll be traveling with my family of four and we’ll be staying from August 6th to August 8th, 2015. My price range is between 150$and$191 per night. The amenities I want the hotel to have are breakfast, a/c, and a gym | 9/495 = 1.81% |
| (mom or mother or dad or father or husband or son or daughter of sister or brother or child etc) | My husband and I would like to take a trip to Wild Rose, Wisconsin from Sept 28th to October 2nd, 2015. We’re looking to stay in an apartment and are willing to pay at least $239 per night. Please make sure this apartment is pet friendly, has a pool, is air conditioned and has laundry and kitchen facilities available | 11/495 = 2.22% |
| sleeps () | I am searching for an apartment that sleeps four in Nevada City, California for the dates of August 28, 2015 through September 7, 2015. The cost should be less than $288 per night and must include TV and wifi. | 2/495 = 0.40% |
| not specified or not found | I’m looking for a hotel in Wilmington Manor, Delaware that is less than 132 per night from 6 Aug, 2015 to 11 Aug, 2015 | 70/495 = 14.14% |

Table 3.2: Patterns for the Number of People

While not necessary for completing a search, the number of people field is necessary for checking accommodation availability. This was done by analyzing patterns in the training data. The patterns used and the number of times they were matched in the training data can be found in Table 3.2. The most frequent (73.3%) query was of the pattern “A hotel room for 3 people" where "people" could be switched out for "persons" or "guests". Other phrasings included "me and 2 family members" or, "for 2 of us." These both occurred 3.23% of the time in the training data. To take into account searches where the user explicitly mentions the other occupants, such as in the search "a hotel room for me, my brother, and my mom," occurrences of these keywords (brother, mom, sister, etc) were counted. The default, in the cases where either the number of people was not provided or not captured, was set to 1.

### 3.2.5 Price

| Pattern | Example | Percent Matched |
|---|---|---|
| ((less or cheaper) than)) or (at most) or (under) or (below) | We are searching for a hotel room in Valley Falls, South Carolina for 4 people. Our budget is less than $205 per night. Our stay is scheduled for 13 Jun, 2015 to 15 Jun, 2015. We would really like a place that offers a pool, a gym, parking and TV. | 143/495 = 28.89% |
| ((more or greater) than)) or (at least) or (over) or (above) | My husband and I would like to take a trip to Wild Rose, Wisconsin from Sept 28th to October 2nd, 2015. We're looking to stay in an apartment and are willing to pay at least $239 per night. Please make sure this apartment is pet friendly, has a pool, is air conditioned and has laundry and kitchen facilities available | 130/495 = 26.26% |
| (between or range of) (and or to or -) | I am looking for a hotel for 3 people in Breckinridge County, Kentucky. I am willing to spend between 178-223 per night, and require the room from the 18th to 19th of April 2015. The hotel should be kid-friendly and have a pool | 86/495 = 17.37% |
| ($) per night | I am looking for an apartment for 3 people near Fort Sumter, South Carolina from May 16, 2015 to May 23, 2015. Ideally the apartment would be pet-friendly, have a pool, air conditioning, and designated parking. My price range starts at $136 per night | 77/495 = 15.56% |
| (number) | I am looking to book a hotel with parking in Bird Island, Minnesota from Feb. 2, 2015- Feb. 6, 2015. There will be three people in the room and I would like to pay 183−241 a night | 17/495 = 3.43% |
| not specified or not found | I need a hotel room in Newtonsville, Ohio for 4 people. The dates are October 12 to October 17, 2015. Room must be pet-friendly, kid-friendly with a-c and a TV | 42/495 = 8.48% |

Table 3.3: Patterns for the Price Range

The price range was likewise determined by patterns in the natural language sentences from the training set. See Table 3.3 for the patterns and the number of times they matched in the training data. The majority of searches contained phrases like "must cost less than $200." There was variety in the use of "less than" or "greater than", and "cheaper" or "greater", and "below" or "above", but these phrasings accounted for 55.2% of matches in the training data. The occurrences of "greater than" or other similar phrases likely stem from the fields provided in the AMT task for the data. The specification for the price in the the task was worded as "Less than $200" or "Between $100 and $200" so it is possible that those phrasings were used exactly as they were seen in the task. Other patterns included ranges, such as “must cost between \$100 and \$200,” which accounted for 17% of the results in the training set. A match was also created for numbers with dollar signs. Since this was not a required search parameter, no default value was assigned.

### 3.2.6 Rating

| Pattern | Example | Percent Matched |
|---|---|---|
| ((at least) or (more than) or (minumum of)) () star | I’m looking for a hotel room for myself in Spring Valley, Nevada. By budget is between 288 and 357 per night. I would like the hotel to be at least a two star hotel that is pet-friendly and have a gym and air conditioning. I want to check in on the April 4, 2015 and check out on April 7, 2015 | 55/112 = 49.11% |
| () stars | I am looking for hotel space for 1 person in Grady County, Oklahoma. It should cost less than \$230 per night. It is needed between 25 Oct to 26 Oct 2015. It should provide breakfast, laundry,kitchen, and wifi, and should be 3 stars | 34/112 = 30.36% |
| ()-star | My family of four will be traveling to McLean County, Illinois from July 18th to July 23rd, 2015. We are looking for a 3-star or higher hotel that is pet-friendly, has adequate parking and complimentary breakfast. Our budget will allow a maximum of \$226 per night | 20/112 = 17.86% |
| not found | | 3/112 = 2.69% |

Table 3.4: Patterns for the Rating Range

A third value was extracted to help filter the results. The model searched for occurrences of star ratings in the original query. This field was not provided in all instances of the data, in fact only 112 of the original 495 queries included references to a rating value. Table 3.4 shows the patterns searched for. Of these, 49.1% were of the form “must have at least 3 stars.” An additional 48.2% were of the form “3 star” or “3-star.”

## 3.3 Accommodation APIs

In order to acquire actual information on accommodation listings, APIs from the Expedia Affiliate Network (EAN) and Zilyo were used. See Appendix A for a complete list of sites that Zilyo accesses. Each of these were called using, at the very least, the location and dates. EAN also allowed the search to be restricted by the price minimum and maximum, the maximum and minimum rate, and the number of people. If found in the search, these fields would also be included in the call to EAN. Zilyo only allowed for searches using a location, number of people, and date range.

Results from EAN included a list of hotels, sorted by "overall value" [1]. For each hotel, the result included the name, a description, rating, and available amenities. Zilyo's results returned an accommodation type (as they have listings for several different providers for various accommodation types) as well as a description and a rating. If the type specified by the user was a hotel, EAN would be used for results. If it was for a hostel, apartment, home, etc, Zilyo was used. In the case that no type was specified or found in the original search, both resources were queried.

## 3.4 Accommodation Results

Results from both Zilyo and EAN were returned in an order ranked by the respective API. These were used as a base ranking. From there, the overall ranking was adjusted based on two metrics: the availability of amenities requested by the user, and the similarity between the search and the result description.

Both EAN and Zilyo return a list of amenities for each accommodation. In order to take into account the user preferences, it is necessary to filter out results that cannot accommodate the user's request. To do this, a list of amenities was collected from Expedia's database (see Appendix B). Variations on these strings were added to the list, to accommodate phrasing differences. An accommodation bigram score was given to each result, comparing the amenities requested and the amenities available. These amenities were turned into bigrams, and used to match against the user's search and the accommodation's results. Bigrams were used in order to accommodate for more specific requests, for example, if the user wanted a "complimentary breakfast", it would not match on accommodations that served breakfast if it did not specify complementary. For amenities that are often only single-word phrases

[1] The "overall value" is calculated by Expedia

(such as “pool”), corresponding bigrams were added to the list. If the score was 0, indicating that none of the requested amenities were available, the accommodation result was discarded. Otherwise, the score was factored into the base score given by the API.

The second metric was calculated to accommodate for aspects indicated in the original search but not covered by one of the aforementioned fields. It was calculated by computing the cosine similarity between the natural language search text and the accommodation's. This was multiplied by a factor of n in order to have a comparative effect to the overall score. From experiments, we found n = 20 to be the best value that would allow these scores to have an effect in the final results.

To produce the final ranking, these two scores were used along with the original rank given by Expedia or Zilyo. Because Expedia and Zilyo both have more experience with returning “best” results for a given search, we did not modify their ranking too heavily. Instead, the modified cosine similarity and the accommodation scores were used in conjunction with the original ranking in order to take into account aspects of the search that were not used to query. For example, if the natural language search includes the phrase “it should have a great view of the beach”, and an accommodation's description also has that phrase, that accommodation will be ranked higher. To update the original rank with these two new scores, the modified cosine similarity and the accommodation bigram scores were subtracted from the original ranking, which was then displayed in increasing value. For example, if a result was originally listed as the 3rd best match, it would have an original score of 3. If additionally it had 2 out of 3 requested amenities (and therefore and accommodation score of 2), it would now have a score of 1. From there, the cosine similarity between the descriptions is factored into the score, potentially moving it “higher” in the result list. It is possible that a result has a negative match score because of the subtraction. As such, matchings with a higher similarity will be ranked higher.

## 3.5 User Interface

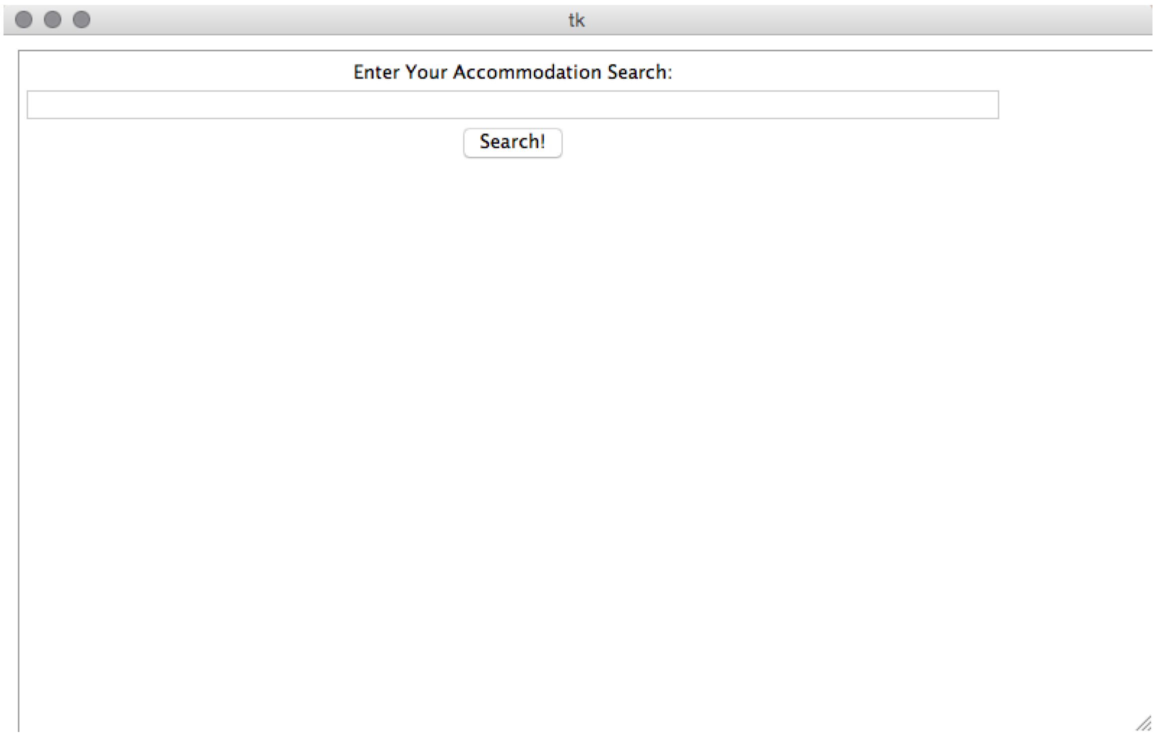


Figure 3.3: UI Home Page

In order to make our system accessible, a user interface was created. See Figure 3.3 for an image of the home page. The user can submit a natural language query by typing in the text field and clicking the "Search!" button. Along with the results, each extracted field is displayed in order to verify that it is accurate. Next to each extracted field is a search area for that specific field that can accept natural language queries. This allows the user to able to change incorrectly extracted fields without having to retype the entire field. If, for example, the user searches for a hotel in Los Angeles for 2 people, but then decides to search for 3 people, that field can be updated directly without the need to update the original query. The results update in correspondence to the updated search. See Figure 3.4

for a visualization of the results and update method.

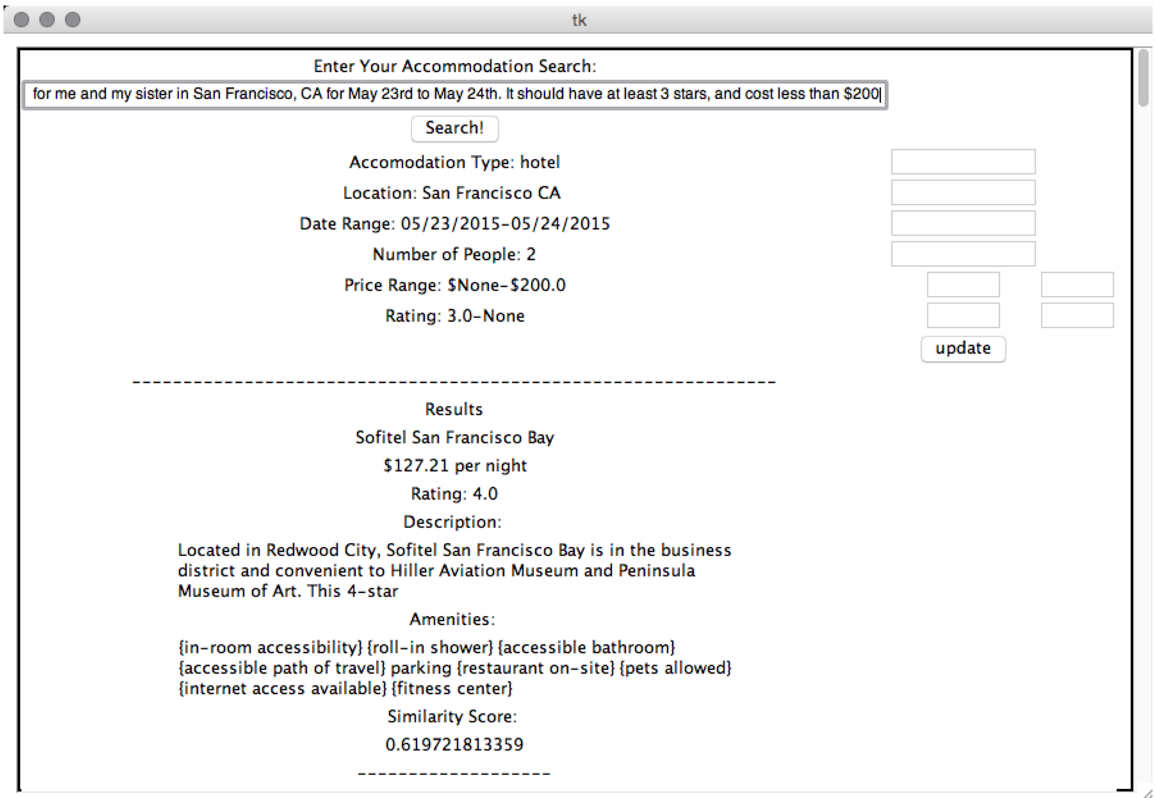


Figure 3.4: UI Result Page

# Chapter 4

# Results

This project has three distinct aspects to it: the field extractions, the accommodation results, and the use of a natural language interface. The field extractions required creating accurate models with which to identify specific fields in the natural language search that were used to query the APIs. With the results, the next step was then to incorporate other elements from the natural language search in order to return more relevant results. Finally, a user interface was designed to accommodate for these types of searches.

Results for the extractions were calculated manually, evaluating how frequently the field extraction models correctly identified the user's request. The accommodation results were taken directly from rankings established by the independent APIs. Without access to how the rankings were calculated, and because the ranking system was not the main focus of this project, these were simply updated in order to accommodate the amenities listed in the search and the similarity between the description and the results. This however, resulted in only a slight modification to the rankings. Because of this, the accommodation results were not separately evaluated. Finally, in order to evaluate the user's preference for a natural language search engine as opposed as to a standard search engine, a user study was performed.

## 4.1 Extractions

| | True | False | Partial | N/A |
|---|---|---|---|---|
| Accommodation Type | 110 | 0 | 0 | 0 |
| Location | 80 | 13 | 17 | 0 |
| Dates | 92 | 2 | 13 | 3 |
| People | 90 | 18 | 1 | 1 |
| Price | 73 | 18 | 17 | 2 |
| Rating | 14 | 1 | 4 | 91 |

Table 4.1: Extraction Results

The extraction process was tested on 110 held-out natural language accommodation searches, acquired through the AMT task described previously. Fields were extracted from these searches and manually verified. Each field was marked as either true, false, partially true, or not applicable. True fields are those that were completely and correctly identified by the extraction model. Partially true fields were marked if the extraction was only partially correct. For location, this was if the city or county was correctly identified but the state was not, or vice versa. In terms of date, these instances occurred when only one of two dates in the range was correctly identified. The same principle applies for price ranges and rate ranges. A field was marked as "not applicable (N/A)" if the value for the field was not specified in the search. This was often the case for a rating, as it was not requested for the entirety of the dataset.

Table 4.1 outlines the results on these 110 natural language searches. The accommodation type was most successfully extracted, with 100% of the type being correctly identified within the dataset. The date extraction was the second most accurate, with 87.37% accurately extracted, and 95.45% were at least partially correctly identified. Identifying the number of people was almost as accurate, correctly extracting 81.81% of the people specified in the training data. In terms of location extractions, the Stanford NER was used (see

System Description section). It successfully identified 72.72% of the locations specified in the searches, and up to 88.18% at least partially correct. The price was at least partially extracted for 83.33% of the testing data. The rate range was correctly identified for 73.68% of the data it applied to.

Each of these models performs sufficiently well to form the basis of the natural language search engine.

## 4.2 User Experience

| | Natural Language Interface | Expedia's Web Interface |
|---|---|---|
| Easiness | 6/10 | 4/10 |
| Speed | 4/10 | 6/10 |
| Naturalness | 8/10 | 2/10 |

Table 4.2: User Study Results

In order to evaluate the preference for natural language search over traditional search engines, a 10-person user study was conducted. For the study, each participant was asked to formulate an accommodation search, including the type of accommodation, the location, the date, the number of people, and any other desired aspects. They were then asked to carry out the search using first the natural language interface designed for this project, and then Expedia's online search interface (available at expedia.com). Given that the hotel data comes from Expedia, the users were advised that the ranked results for both the natural language interface and for Expedia's would be the same and that was not the focus of the study. They were then asked to choose which they preferred in the following categories: which was easier to use, which was faster to use, and which was more natural to use. Table 4.2 provides an overview of user's preferences towards using the natural language interface versus Expedia's interface in each of the three categories.

### 4.2.1 Easiness

Six out of ten users specified that the natural language search engine was easier to use. Comments included that it was easier to put all of the information in one place, and that the result page was also helpful, as it allowed users to update specific fields instead of the entire natural language search.

### 4.2.2 Speed

Users indicated that Expedia's interface was faster 6 out of 10 times. However, one commented that "In reality they were about the same," and that factoring in the time to filter out amenities could make the natural language interface easier. "Speed" was not defined, and some users interpreted it as speed of returning results while others took it to mean the speed of entering the search.

### 4.2.3 Naturalness

The third aspect that the users were asked to compare was which system was more natural to use. In this respect, users chose the natural language search 8 out of 10 times. The comments for this section provided the most insight into user's preferences towards a natural language interface for these types of searches. While some users found it more natural to type a search the way they constructed it in their mind, others thought that typing this out was less natural because it wasn't what they were used to and were less familiar with constructing a search in that way and preferred to simply fill in the predefined fields.

# Chapter 5

# Conclusion

The increase in popularity and desire of users to arrange their own travel has created the demand for online interfaces for travel products. An integral subsection of the travel industry is accommodations, and with the rise of sites like AirBNB and CouchSurfing, consumers are becoming more and more creative in their demand for these products. The interfaces available to make queries for these resources has not kept up with this demand. In this project, we investigated how well a natural language interface performs for accommodation searches.

Extracting fields from a natural language search proved to be successful, allowing users to formulate requests in whatever way proved most intuitive to them. We were successfully able to create models that identified and extracted each of the fields required to search the pre-existing databases (Expedia and Zilyo). With an average success rate of 76% and a method for updating the results for individual fields, this proved to be suitable for the overall performance of the system and proved that identifying these fields in natural language is a feasible task.

In terms of user experience, this project leaves room to improve. It is difficult to distinguish between users who truly find that a natural language interface is less intuitive, and those who are simply not used to having that option. It is possibly a by-product of the generation of young adults who have grown up conducting their own travel queries online, trained to formulate searches in a specific manner. However, the users proved to be receptive to the concept which indicates that, should natural language interfaces to this data become more widespread, this could be a successful approach.

As a result, the expansion of natural language search engines in this field and their ability to capture the entirety of an accommodation request will better demonstrate the usability and preference for interfaces of this type. This project has indicated that there is a space, and a need, for products of this kind and that there is room to improve in the field.

# Chapter 6

# Future Work

While this project outlines the approach for a basic natural language interface for accommodation searches, opportunities remain to improve this process and the impact it could have on the field of accommodation searches, and more broadly, travel searches.

Data for this project was collected using Amazon Mechanical Turk tasks with a predetermined set of parameters. This biased the data used to train the field extraction models. To create models that are capable of parsing a greater variety of natural language inputs, a larger and more varied set of data would need to be trained on. However, the limitations on acquiring this sort of data, outlined in the previous section, have not changed and likely will not change until natural language interfaces are more widely available.

Given the data the field extraction models were trained on, they do reasonably well to accurately recognize these fields. However, we could improve these models with more complex natural language processing tools. This project uses a series of regular expressions to match patterns in the text and as such, they only recognize a specific set of phrase types. Improvements to this would add the ability to recognize a greater variety of phrase patterns. This would help with recognizing incorrect spelling as well as improving the date and price extractions in particular.

In terms of data retrieval, to improve this project we would benefit from access to more databases for accommodation results. This project only uses data from Expedia and Zilyo, which limit the types of accommodations available and the information available for each accommodation. With additional information on individual listings, we can do a better job of matching a search to a result because more information about the result would be available.

While the focus of this project was not the user interface, for future work we hope to improve the design and functionality of the UI.

Finally, the user study we conducted was limited by the demographics of the users and their experience with the internet. A more robust user study could reveal further results on the use of a natural language interface.

Ultimately, this project defines an approach for creating a natural language interface for accommodation searches, proving that not only can one be created with a high level of success and user-satisfaction, but also shows the promise for a natural language interface in this field.

# Appendices

# Appendix A

# Sites Accessed by Zilyo

1. 9Flats.com - www.9flats.com
2. airbnb - www.airbnb.com
3. AlwaysOnVacation - www.alwaysonvacation.com
4. Apartments Apart - www.apartmentsapart.com
5. BedyCasa - www.bedycasa.com
6. Cities Reference - www.citiesreference.com
7. Geronimo! - www.geronimo.com
8. gloveler - gloveler.com
9. Holiday Velvet - www.holiday-velvet.com
10. BookingPal - holidayvillas.com
11. HomeAway - www.homeaway.com
12. Homestay.com - www.homestay.com

13. Hostelworld.com - www.hostelworld.com

14. House Trip - www.housetrip.com

15. Interhome - www.interhome.com

16. roomorama - www.roomorama.com

17. The Other Home - www.theotherhome.com

18. travelmob - www.travelmob.com

19. vaycayhero - www.vaycayhero.com

20. Way to Stay - www.waytostay.com

21. WebChalet - www.webchalet.com

# Appendix B

# Amenities Used

1. Business Center
2. Fitness Center
3. Hot Tub On-site
4. Internet Access Available
5. Kids' Activities
6. Kitchen or Kitchenette
7. Pets Allowed
8. Pool
9. Restaurant On-site
10. Spa On-site
11. Whirlpool Bath Available
12. Breakfast

13. Babysitting
14. Jacuzzi
15. Parking
16. Room Service
17. Accessible Path of Travel
18. Accessible Bathroom
19. Roll-in Shower
20. Handicapped Parking
21. In-room Accessibility
22. Accessibility Equipment for the Deaf
23. Braille or Raised Signage
24. Free Airport Shuttle
25. Indoor Pool
26. Outdoor Pool
27. Extended Parking
28. Free Parking

# Bibliography


[BHCDV99] L. Braden-Harder, S.H. Corston, W.B. Dolan, and L.H. Vanderwende. Apparatus and methods for an information retrieval system that employs natural language processing of search results to improve overall precision, February 4 1999. WO Patent App. PCT/US1998/009,711.

[JS06] Bernard J. Jansen and Amanda Spink. How are we searching the world wide web?: A comparison of nine search engine transaction logs. *Inf. Process. Manage.*, 42(1):248–263, January 2006.

[Red98] D.M. Redfern. Natural language meta-search system and method, June 18 1998. WO Patent App. PCT/CA1997/000,970.

[Str95] Tomek Strzalkowski. Natural language information retrieval. *Information Processing and Management*, 31(3):397 – 417, 1995. The Second Text Retrieval Conference (TREC-2).

[Tur94] Howard Turtle. Natural language vs. boolean query evaluation: A comparison of retrieval performance. In *Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '94, pages 212–220, New York, NY, USA, 1994. Springer-Verlag New York, Inc.

[XP11] Zheng Xiang and Bing Pan. Travel queries on cities in the united states: Implications for search engine marketing for tourist destinations. *Tourism Management*, 32(1):88 – 97, 2011.