When To Go Where

Samyuth Sagi - sagis2@rpi.edu
Tyler Layton - laytot@rpi.edu
Annabelle Choi - choia7@rpi.edu
Ben Rodgers - rodgeb@rpi.edu

Abstract
Choosing the right national park to visit can be challenging due to factors such as temperature, trails, and crowd levels, which can all significantly impact the experience. With the variety of environments and conditions found across national parks, adventurers often struggle to find a park that suits their preferences. This project aims to develop a recommendation system that helps users select the ideal national park based on specific preferences, such as preferred temperature, trail type, and less crowded times. Integrating data into the system will provide tailored recommendations that simplify the park selection process. This tool is intended to make it easier for adventurers to find a park that meets their needs, ultimately enhancing their outdoor experiences.

1 Introduction

In 2023, over 312 million people visited U.S. national parks, leading to overcrowding during peak seasons. This high level of visitation makes it challenging for visitors to choose the right park, as factors like weather, trails, and crowd levels can significantly impact their experience. Many visitors, especially first-timers, struggle to find a park that suits their personal preferences. This project aims to develop an ontology-based recommendation system to help users select the ideal national park based on factors such as temperature, terrain, and crowd levels. By integrating data on weather patterns, terrain types, and visitation trends, the system will offer tailored park suggestions, addressing the needs of both experienced adventurers and newcomers.

2 Technical Approach

Preliminary Work

The development of the national park recommendation system started by gathering data on park visitation, temperature, and crowd levels. We created an ontology to organize and define concepts like ”temperature,” ”trails,” and ”crowd levels”.

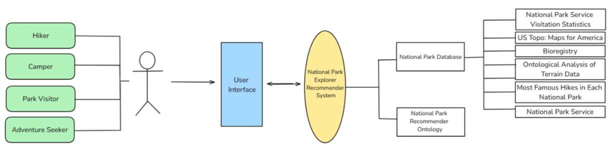

Figure 1: System Diagram

Conceptual Model

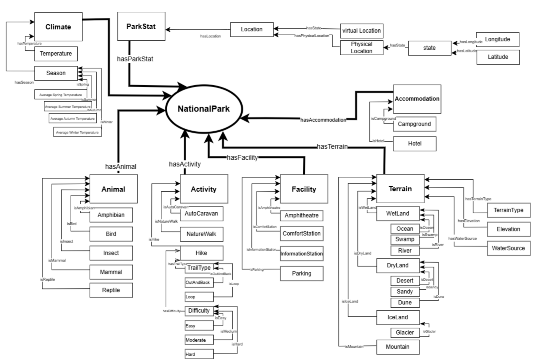

Figure 2: Conceptual model of the recommendation system.

Ontology

The conceptual model above is a visual representation of the ontology that we developed in this project (Figure 2). In the above we can see that there are a set of concepts such as Animal, Activity, Accommodation, Facility, Park Statistics, Terrain, and Climate. All of these concepts are attached to a central concept National Park which is the focus of our ontology. They also have sets of other concepts that are connected to them such as Accommodation which is attached to Campground, Hotel, Capacity, Availability, and Amenities.

The conceptual model also shows how we have connected the different concepts in our ontology. Using the relations we have defined we are able to query our ontology to answer specific questions. For example, to find the longest trail in California, the system would look at parks in California, check which trails they have, and find the one with the greatest length. The conceptual model helps the system understand and process these connections efficiently.

Data Scraping

With the amount of information that we wanted to include in our ontology, we needed an easy way to get data from the internet about the National Parks. To pull data we found information that was publicly available from a number of sources with the primary source being the National Park service. We were able to obtain some data from manually annotating information available on the website, pulling data reports, web scraping, and accessing public APIs. The information that we sourced in this project focused on animals available in each National Park, temperature and visitation statistics, locations, and hiking trails.

3 Related Work

Gathering a multicultural ontology of outdoor and hiking landmarks from direct terrain observations and sketch (found here) *maps* is an ontology that can be compared to ours in some senses. The main difference is this ontology is used for European countries, so it has no application at all to our specific National Park ontology. The comparison here lies in the fact that they care about where the hiking landmarks are and the types of trails present. This ontology is much simpler than ours, as its concept map is much smaller, but it is related to our work in the sense that it is outdoor/adventure-related.

4 Evaluation

Competency Questions

Competency Question 1: "If I want the coldest temperature park in the midwest during summer, what park should I go to that fits my comfort level?"

The system queries the ontology to identify all parks located in the Midwest region using the location attribute. For these parks, it retrieves temperature data for the summer months (e.g., June–August). It then filters the results to find the park with the lowest average summer temperature, ensuring it matches the user's comfort level (in this case, a peak summer temperature of 70°F). In the usage scenario, John seeks an adventurous hiking experience in the Midwest. The system provides Rocky Mountain National Park as the recommendation, as it has the lowest summer temperature among Midwest parks, meeting his criteria. The decision relies on analyzing location, temperature, and park data stored in the ontology.

Competency Question 2: *"I am new to hiking. Which national park has cool summer temperatures and hikes less than 2 miles?"*

To answer this competency question, the system begins by querying the ontology to identify all parks and their temperature data for summer. It filters parks with cool temperatures (e.g., around 65°F). Next, the system accesses hike data for each park and retrieves trails shorter than 2 miles. By combining these criteria, the system identifies Crater Lake National Park, where summer temperatures average 65°F, and the Watchman Peak trail is 1.7 miles, making it suitable for a beginner like Jeb. In the usage scenario, Jeb's specific preferences for a short trail, cool temperatures, and a scenic location for photos are satisfied by this recommendation. The process uses data from the season, temperature, park, and hike attributes in the ontology to deliver an accurate and tailored result.

Competency Question 3: *"I am in California for a week and I am curious as to what the longest hikes out of all the National Parks here are."*

To answer this competency question, the system queries the ontology to identify all parks located in California using the location attribute. It then retrieves data about the trails within these parks, compiling a list of hikes and their distances. From this list, the system determines the longest hike, identifying the Lakes Trail as the result. In the usage scenario, Jeb's curiosity about the longest trail in California aligns with this search, as he seeks a scenic adventure that fits his preferences. This process uses location, park, and trail attributes to efficiently deliver the information based on the user's query.

Competency Question 4: *"I am near Acadia and I want to see an American bison, is the American bison native to Acadia National Park?"*

This competency question checks whether the American bison is native to Acadia National Park by querying the ontology for wildlife associations. The ontology defines relationships between parks and the animals that inhabit them. The query reveals animals found in Acadia, such as the Black Fox, Coyote, and Moose, as a backup as they are other animals in the park. This information can help visitors set realistic wildlife expectations and guide park management in wildlife conservation planning.

Competency Question 5: *"If I want to go to the northernmost park in the United States that is the least visited in Winter, where should I go?"*

The system first uses ontology to identify the northernmost parks in the United States by querying the location attribute. Among these, it filters parks based on visitation statistics during the winter months, identifying the one with the lowest visitor numbers. Additional factors, such as weather conditions, may be considered to ensure winter accessibility. In the usage scenario, Joe inputs his desire to see a moose, avoid crowds, and explore northern parks. The system compares park data, including wildlife presence, visitation rates, and hiking opportunities, to recommend Arctic Gates National Park as the northernmost and least visited park in winter. This conclusion is derived by analyzing data stored in the ontology and matching it to Joe's preferences.

5 Discussion

Value of Semantics

In our ontology, semantics played a large role, especially when we were creating it. One of the first focuses is how the National Park needed to be clearly defined. This was executed well within our project, as we classified a 'National Park' as the items in the set of National Parks designated by the US government. A second focus was on how we wanted to define a facility. In our ontology, we have both accommodations and facilities. To be specific and clear, we need to determine what makes a facility a facility. This was defined as a facility being someplace you can use to your advantage, but not improve your sleeping situation. For example, a comfort station is not the same as a campsite. We defined a campsite as an accommodation because you get to benefit by sleeping there, meanwhile, a comfort station is a facility because you are not using it to sleep. The semantics we chose here provide a definition of how a visitor can use these to their advantage. One of the other focuses was on location and how we chose to classify it. This might seem like an easy play at first. We should just do the state of the park and leave it at that. This is very intuitive, though with more thought, it comes with its own set of issues. For example, what happens when there are multiple parks in one state and you are on the opposite side of the state of which your recommended park is in? So in our semantics, we didn't label one park with "location". Each park has a property of hasState, which contains the state it is in, but they also have their longitude and latitude attached as well. This allows the user to specify how far north, south, east, or west they are willing to travel and opens up the ontology for more specific queries. If we had more time to work on this project in future semesters, we would add more to our semantics to allow us to do more inference. A certain example of this would require taking advantage of the data we already have. Assuming we implement a travel planning and state park aspect to our ontology, we could make an inference using visitation statistics of the National Parks in that state. For example, if we know that Gates of the Arctic and Denali are low in visitors relative to other National Parks, we can infer that this will affect ticket prices in a certain manner and that the state parks in that state might also experience low visitation. This way we can plan a solid trip through inference. Using the statistics we collected already, it would be simple to implement a way to assume more about the attractions close to that park or the state in which the park resides.

Link to Project Site

Link to When-To-Go-Where project site.

Limitations and Scope

One of the biggest limitations of our current work is the data that we were able to gather. When attempting to make an ontology for all 63 national parks with a way to display them to a hopefully traveler, this is the first issue one will come across. There are slight things that come up later than you would think. For example, you want to find out if a park is in a certain state. Just from this task, you have to gather the states of all the parks, then define a state in your ontology, place the parks in those states on the ontology level, all before this can be displayed to the user. That was just a simple example. We wanted to incorporate animals in our ontology, as this would be good for the animal loving/hunting demographic. The problem with this is that there are far too many animals to account for in the United States, after all, it is an extremely biodiverse country. If there were prebuilt software to gather and place this data, I am sure we could have made more progress this semester, but there isn't. We ended up scoping this down and included only a few animals for the most popular parks.

It is also important to note that in our conceptual model, there are a lot of areas which we don't use/don't have individuals for. This is because we could find the data or we didn't have enough time to get the data. We wanted to have our full ontology, but we realized too late that this was going to be too hard so we had to scope the project down. The classes are still there in the ontology for future work if more data is found.

Future Works

Our current scope is restricted to recommendations for the 63 United States National Parks. Some potential expansions include expanding the number of parks in the system.

We could for example extend our project to state parks however this is tricky as we have not been able to identify as many consistent sources of data for these parks. Supposing the data issue was not present, however, the approach would fundamentally be the same as our approach to making recommendations for all of the national parks. We would not need to even make any changes to our current conceptual model. Another extension that could be considered is extending to National Monuments however this would require us to make larger changes to the current conceptual model.

Another direction we can expand our project is to lean into the planning aspect of the park. Individuals might not be as clued in as they think they are when it comes to preparing for these parks. For example, in the Gates of The Arctic National Park, there are absolutely no amenities. If you want to go to an official campground, you can't. This is a stark difference when compared to parks like Yellowstone which have glamping sites for anyone wishing to spend the money on it. If the individual doesn't do specific research on the gear they need or aren't provided with the steps to take to prepare, they might not be in the best position for their trip. Much like in the wine ontology, this next step in our ontology could provide the user with links and lists to help them prepare.

6 Conclusion

Our idea of creating an ontology-based recommendation system to help users select the ideal national park based on factors they specify was executed to the best of our abilities. We modeled our system starting with the overall class of what defines a National Park and moved down the chain of importance into classes like weather, trails, and visitation statistics. Once we have the weather class, we are able to find more descriptive words like climate. Through climate, we can reach descriptors like hot or cold efficiently. This model allows the user's query to flow efficiently to each channel it needs to go through and bring back the proper selection. When it comes to the competency questions we started with, our model performs relatively well compared to when we first started asking it questions.

7 Acknowledgments

We would like to thank the following individuals for their guidance and support throughout this course:

Prof. Deborah L. McGuinness and Ms. Elisa Kendall, for their expertise and instruction.

Jade Franklin, Vladia Pinheiro, Kelsey Rook, and Danielle Villa, for their dedicated coordination and assistance.

References

[1] National Weather Service, "Glossary of Weather Terms," Weather.gov. [Online]. Available: https://www.weather.gov/cae/weatherterms.html. [Accessed: 25-Sep-2024].

[2] Merriam-Webster, "Merriam-Webster Dictionary." [Online]. Available: https://www.merriam-webster.com. [Accessed: 25-Sep-2024].

[3] FATMAP, "The Most Famous Hike in Every US National Park," FATMAP Guidebook. [Online]. Available: https://fatmap.com/discover/guidebook/united-states-of-america/washington/alpine-climbing/hiking/the-most-famous-hike-in-every-us-national-park/109220. [Accessed: 23-Sep-2024].

[4] Integrated Taxonomic Information System, "Integrated Taxonomic Information System." [Online]. Available: https://www.itis.gov/. [Accessed: 24-Sep-2024].

[5] Encyclopedia of Life, “Species Profiles.” [Online]. Available: https://eol.org. [Accessed: 25-Sep-2024].

[6] National Park Service, “Wildlife in National Parks.” [Online]. Available: https://www.nps.gov. [Accessed: 25-Sep-2024].

[7] National Audubon Society, “Bird Guide.” [Online]. Available: https://www.audubon.org/bird-guide. [Accessed: 25-Sep-2024].

[8] Smithsonian Institution, “Animal Information and Conservation Resources.” [Online]. Available: https://www.si.edu. [Accessed: 25-Sep-2024].

[9] National Park Service, “Species List (IRMA Portal).” [Online]. Available: https://irma.nps.gov/NPSpecies/Search/SpeciesList. [Accessed: 25-Sep-2024].

[10] National Park Service, “Glossary of National Park Service Terms.” [Online]. Available: https://npshistory.com/publications/nps-glossary.pdf. [Accessed: 25-Sep-2024].

[11] P. Kettunen, E. Nuhn, M.-D. Van Damme, and A.-M. Olteanu-Raimond, ”Gathering a multicultural ontology of outdoor and hiking landmarks from direct terrain observations and sketch maps,” *Abstracts of the International Cartographic Association*, vol. 7, p. 74, 2024, European Cartographic Conference – EuroCarto 2024, 9–11 September 2024, TU Wien, Vienna, Austria. [Online]. Available: https://doi.org/10.5194/ica-abs-7-74-2024.