Semantic Data Models for Hiking Trail Difficulty Assessment

Jean-Paul Calbimonte[1], Simon Martin[1], Davide Calvaresi[1], Nancy Zappelaz[1],
and Alexandre Cotting[1]

Institute of Information Systems
University of Applied Sciences and Arts Western Switzerland (HES-SO)
Sierre, Switzerland
`{firstname.lastname}@hevs.ch`

Abstract. Hiking is a popular outdoor activity that if practiced regularly, can bring significant health benefits. Moreover, considering that hikers range from expert mountaineers to older adults with limited physical capabilities, it touches a large target audience, and is strategically included in several tourism packages across the globe. Thus, a precise characterization of the tracks, especially regarding their points of difficulty, is crucial to effectively cope with the challenge of identifying the best-suited hiking trails for heterogeneous users. This paper introduces a semantic model for representing and integrating the main characteristics of a track, including their different types of difficulties, using Semantic Web ontologies. The construction of knowledge graphs that use such a model may constitute a first step towards a system for personalized recommendations of trails based on difficulty-classification criteria.

Keywords: Hiking tourism · Difficulty assessment· Semantic models · Semantic Web · Tourism Knowledge graph.

1 Introduction

Pedestrian tourism comprises a set of increasingly popular outdoor activities, accessible to people of all ages, physical, and economic conditions [7, 3]. In this sector, the demand for routes and trails is growing drastically. In particular, in countries with a long tradition such as Austria, Italy, and Switzerland, this demand boosted interest in providing more efficient, personalized, accessible, and accurate information (in particular, supporting digital applications). In this context, different complementary actors have been actively contributing to the emergence of both public and private resources, and applications that partially address the information and usage needs of pedestrian tourism. For example, in the region of Vorarlberg[1] in Austria, the local tourism office holds an online database of hiking trails offered to potential visitors. As another example, in the Swiss canton of Valais, the management of official hiking trails is governed by

---

[1] `https://www.vorarlberg.travel`

Valrando[2], offering online descriptions and maps, relying on geographic information sources managed by the federal government. At the same time, SuisseMobile[3] provides online apps that allow discovering and tracking pedestrian trails, offering additional technical information. Complementary to these services, in each location or region, tourism offices provide detailed and curated information about local pathways, including cultural or thematic offers, sometimes through 3rd party exploitation companies such as Outdooractive[4] or Snukr[5]. The collection and maintenance of these information knowledge bases about hiking trails is costly, and requires on-site interventions by local guides, but can also benefit from crowd-based feedback. This detailed information is crucial afterwards, as it provides hikers with essential indications of which trails (or parts of trails) are adapted to their preferences, context, time constraints, and limitations. Given the heterogeneity of these information sources, potential hikers have to deal with complex analysis to choose among existing pathways. Such complexity escalates if hikers have certain limitations, such as fear of heights, climbing constraints, reduced mobility, walking aids, or terrain preferences.

In this paper, we explore the use of semantic data models for representing features of hiking trails, especially those that may constitute an obstacle or a difficulty for specific users. The presence of such features has a direct impact on the user experience. Hence, an "apparently" mild difficulty may completely block a hiker in some circumstances (e.g., a hanging bridge for a person with the fear of heights). Moreover, the objective representation of the path features and the modeling of their perception (users-wise) is not trivial. In particular, besides the intrinsic characteristics of the difficulty points, it depends on the user's context. This work elaborates on the principles of a methodology that considers three main aspects: effort, technique, and risk. Using semantic representations to link these aspects together with the geometry information of a hiking trail, we provide a solid base for the development of location-based services related to hiking recommendations.

The remainder of the paper is structured as follows. Section 2 presents related work. Section 3 presents the main principles behind the model. Section 4 explains the methodology chosen for describing the difficulties and information needs. Section 5 describes the semantic model itself, while Section 6 provides a discussion about the potential use of this model and its extension as a knowledge graph for recommendation purposes. Finally, Section 7 concludes the paper.

2   Related Work

Hiking is an outdoor activity that has shown extensive adoption in heterogeneous segments of the population and has proven to be particularly beneficial for peo-

---

[2] https://www.valrando.ch/
[3] https://www.schweizmobil.ch
[4] https://www.outdooractive.com
[5] https://www.snukr.com

ple with limited mobility. The economic interest of maintaining and promoting hiking sites has been shown to increase and gain in terms of market value [16, 15]. Moreover, it has also received the attention of health promotion bodies and institutions, who, over the years, collaborated in improving existing guidelines and making more accessible scientific recommendations [6, 9, 10]. In this context, several initiatives promoting awareness about hiking trail difficulties (e.g., updating tracks conditions according to their situation) arose. For instance, in [12], a chatbot-based approach is described, aiming at providing recommendations of tourism offers. Other efforts, as in [5], considered recommendations for groups of users, taking into account combinations of preferences. Some of the early experiences to assess difficulty in terms of energy expenditure have been reflected in works such as [8]. Other efforts have focused on the analysis of preparedness [11] of hikers, or collaborative annotation of tourism objects[6]. Regarding the suggestion of hiking trails, several works have partially addressed the problem. For example, Boerger et al. [2] and Pitman et al. [14] proposed recommendation algorithms and tailored hiking time estimations, and Calbimonte et al. [4] have tried to use questionnaires to estimate the users' physical condition.

Although some of these works introduced certain elements which are relevant for the evaluation of track difficulties, none of them considers the different aspects included in the model below presented, and do not comply with existing Semantic Web modeling techniques. Concerning standards for representation and modeling of tourism and travel concepts and objects, several initiatives and resources for different scopes and purposes can be acknowledged. For example, the Open Travel Alliance (OTA)[7] has developed standard models for travel objects, such as the Open Travel Schema. This model is restricted mostly to booking information and do not conform to current ontology modeling standards. Another relevant resource is the Thesaurus on Tourism and Leisure Activities published by the WTO (World Tourism Organization) [13], used mainly for indexing concepts in the tourism realm. On the same line, the Travel Technology Initiative[8] provided a set of standards, including different messaging and tourism agent specifications.

Such standards offer given structural guidance and technological means for integration of information in the tourism domain. Nevertheless, they focus mainly on booking operations and availability, and less on the content specified for outdoor activities. In particular, there are no data description models for representing pedestrian roads and the different aspects that analyzed in this paper (e.g., description and specification of difficulty types are entirely neglected by the aforementioned previous works).

3   Hiking trail difficulties

This section elaborates on the challenges underlying the difficulty-based trail characterization of hiking trails. One of the reasons causing such complexity is

---

[6] https://www.apidae-tourisme.com/

[7] https://opentravel.org

[8] https://www.tti.org/

the degree of subjectivity that the difficulty may entail, linked to the relative perception of a trail user. Each person may have entirely different notions of difficulty, associated with their previous experiences, background, fitness level, age, etc. For example, a rocky mountain trailing path may be mildly difficult for a young hiker, but extremely hard for an aged user. Moreover, difficulties can be related to different orthogonal aspects. For instance, a bridge crossing may not be associated with physical difficulty, but if it is located over a deep canyon, it may cause vertigo or other psychological effects on certain hikers. Similarly, a scarped rocky trail may require a specific technique level, although it may not be physically demanding.

In particular, the Swiss hiking offer, reflected in the information available in web sites of local tourism offices and public organizations, includes some information relative to trail difficulties. This information includes basic but essential data, such as the elevation, generally perceived difficulty, total estimated time, and distance (see Figure 1).

Fig. 1. Description of a hiking trail in an online tourism database. The trail includes basic information, including distance and estimated time, but insufficient data concerning the difficulties that it may entail, including obstacles, risks, or detailed physical requirements.

Clearly, these basic information items are not enough to allow people with limited physical, technical or psychological capabilities, to decide which hiking trails are better suited for them. As we have evidenced through questionnaire and workshops conducted with key stakeholders in the Swiss tourism sector [4], this

type of information is highly demanded both for end users and tourism providers, with a high potential for exploitation[9]. From these previous experiences, we identify the following principles:

P1: Self-assessment Given the difference in terms of the perception of the level of difficulty of a certain hiking track segment, enabling every potential user to provide a self-assessment of his/her physical condition is crucial, possibly through specialized questionnaires.

P2: Multi-criteria evaluation Considering that difficulty depends on several factors, often including distinct aspects, any assessment of difficulty levels must include a combined evaluation of these criteria (e.g., combining physical or psychological difficulties).

P3: Effort assessment The required effort for a given hiking trail is a key aspect that must be considered in the evaluation of difficulty. The physical effort can be measured in different ways (e.g., including total distance, and/or slopes).

P4: Technique assessment The technique is another critical aspect to be considered (complementary to the effort). Although a hiking track may require a limited effort, it may demand a certain technical level to be crossed. The assessment of the technique may include evaluating the presence of obstacles, the need for specific equipment, or the use of hands for particular segments of the track.

P5: Risk assessment Risk is often a main discriminant factor for defining the difficulty of a hiking track. Thus, it is another aspect to be considered. It should consider the track geography and geology (e.g., instability, cliffs, and void), as well as other elements increasing the occurrence of an accident (e.g., falls and slipping).

P6: Profile recommendation A model for the difficulty in hiking trails must also include the elaboration of a profile which considers the aspects above, and which can be compared quantitatively with other profiles, so that recommendations can be based on those features. Such a model can be used as a basis for automated and personalized recommendations of hiking trails, which can be proposed to users according to their effort, technique, and risk preferences.

P7: Semantic modeling A difficulty model should provide the necessary abstractions and concepts to describe the different features and aspects detailed in the previous points. Moreover, it should use existing standards and modeling approaches that allow this information to be reused and interlinked with other data sources, and including machine-readable representations. Linked Data and other related Semantic-Web models provide the technical foundations for creating such models, in the form of machine-interpretable ontologies.

---

[9] https://portal.klewel.com/watch/webcast/technoark-2018-quantified-self/talk/10/

4   Difficulty assessment model

To cope with the lack of description and specification of the trials difficulty points characterizing a hiking trail, this work considers the factors mentioned above to propose a semantic model for hiking track difficulties, reusing existing concepts and notions of the hiking and leisure industry. In particular, we have adopted the classification methodology from the French Hiking Federation[10]. Such a methodology aims at providing a simple (yet comprehensive) estimation of the tracks' difficulty (for both experts and novice hikers). Concretely, it defines difficulty according to the following criteria:

- Effort: related to the difficulty associated with the physical energy required for a hiking track.
- Technique: related to the mobility/motion difficulty required for a hiking track.
- Risk: related to the psychological difficulty associated with a hiking trail.

The *effort* is directly related to the physical difficulty, and as such, it can be measured according to the energy required to cross a certain track. Several parameters can be used to estimate the effort. For example, the total distance can be an essential factor. A 2km track will require less effort than a 4km track, if both have similar terrain characteristics. Other factors include the slopes, altitude, descent, and slope changes (see Figure 2). A track with steep slopes may require more energy expenditure than a flat track, even if they comprise the same distance. Although these features of the hiking tracks can be measured, it is not straightforward to come up with indicators that can be directly used for characterizing the difficulty in these terms. A promising initiative in this scope is the IBP index[11], a numerical scale for representing the human effort in hiking and biking tracks. This index is based on an algorithm that takes as input the GPS coordinates of the points that constitute a track, and calculates a numerical value considering the different slope gradients, ascending and descending distances, altitude, etc. Although this is not the only possible way of measuring effort, it is a tool that already has gained recognition from institutions such as the French Hiking Federation, and companies/applications in the field, such as Strava.

The *technical* difficulty of a hiking trail is associated with the motion/motricity required to overcome obstacles present in the track. For example, rocks can be obstacles that require specific skills to be overtaken. Small stones may require raising the feet moderately, while larger rocks may even require the use of one or both hands to overcome the obstacle. Current technologies do not provide yet a (semi)automatic way of estimating this type of difficulty, which is generally assessed by the observation of an expert. In general, an *easy* track has little if no obstacles, not requiring any particular movements or technique, other than a normal gait. More difficult hiking trails may have other types of obstacles, and there could be different ways of classifying them. For example, the French

---

[10] Fédération Française de la Randonnée Pédestre `https://www.ffrandonnee.fr`
[11] `https://www.ibpindex.com`
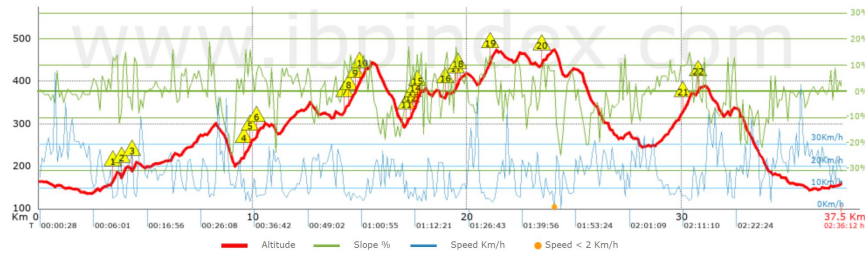
Fig. 2. Analysis of a hiking trail and its different segments, including identification of steep slopes, changes in climb and descent gradients, among other indicators that may relate to a physical difficulty. In this particular case the score is calculated according to the IBP index, although other metrics could be used.

Hiking Federation uses the type and height of feet movements required to cross through an obstacle, as a reference for difficulty assessment. An obstacle requiring a movement up to the height of the ankle has significantly less technical difficulty than a movement up to the height of the knee (or even more if to the height of the hip or demanding to use the hands or walking sticks).

Regarding the *risk*, it entails a psychological difficulty, related to potential accidents or situations to which the hiker is exposed. As in the previous case, it persists the difficulty to provide an assessment of risk, and in general, it is a field expert who determines it in accordance to given guidelines. As an example, urban and rural itineraries without remarkable terrain irregularities can be considered as implying low risk. Even in case of a fall, in such circumstances, the consequences would be generally benign, as opposed to other terrains (e.g., with pronounced slopes and/or very irregular and narrow paths) where a fall could even cause fractures or other more severe injuries. Moreover, in high-risk tracks, such as mountain trails, a fall could even have fatal and nearly-fatal consequences (e.g., in the presence of high cliffs and abysses).

As we have seen, in the cases of risk and technical difficulty, specific obstacles or terrain characteristics need to be clearly profiled to perform an accurate assessment. It is worth noticing that in some cases the presence of a single obstacle may be enough to change the classification of a track. For example, an otherwise entirely plain track could have a single point in which a high cliff would force it to be considered as high-risk. Hence, for hikers with limitations (e.g., fear of heights/depths) this single point of difficulty would be enough to block the entire trip.

Having described the main notions of difficulty, the model described in this paper needs to answer to the following competency questions, which represent the information requirements.

– What hiking tracks are available within a given region?
– What are the coordinates of a particular hiking track?

- What is the total distance of a given hiking track?
- What is the accumulated ascents and descents of a given track?
- Which interest points are near my current location in this track?
- Which tracks require a mild physical effort according to my capacity?
- Which tracks require advanced technical skills?
- Which segments of a given track have steep slopes?
- Which tracks have a low-risk difficulty and high effort difficulty?
- Which tracks are similar to a particular hiking trail in terms of overall difficulty?

5   Semantic model for difficulty assessment

As detailed in the design principles of this model, we have chosen to follow a representation that reflects the concepts in this topic, and it is aligned with current trends in semantic data management. The chosen semantic model follows the principles of Linked Data [1], and uses the Resource Description Framework (RDF) as a basis for its technical implementation. The resulting *ontology* formalizes the domain concepts and is machine-readable. Thus, it can be fed to reasoners and other semantic data engines. Furthermore, it can be used to build a hiking knowledge graph, potentially exploited by querying and analytics tools.

The first step for this model is to represent the hiking trails. It includes name, description, geographical coordinates, places, points that are part of such an itinerary, and ratings and reviews. As shown in Figure 3, we have followed best practices in ontology engineering, in this case, by reusing a widely used model such as schema.org[12]. This vocabulary has gained adoption in the last years, becoming one of the leading general-purpose ontologies that feed knowledge bases on the Web.

An essential portion of the information needs is already present in schema.org, including geographical emplacement and basic information about the elements that compose the track, which can be represented as a special type of *Tourist-Trip*. These are represented as *Places*, which could be either points, segments, or other geometries. It is possible to use the geographical information to perform spatial operations, such as calculating distance and proximity, or containment of tracks within tracks, or points of difficulty within a specific hiking trail. Using schema.org provides the additional advantage of increasing the potential compatibility with other tourism data sources published on the Web, overcoming heterogeneity issues.

Nevertheless, the data defined by schema.org for tourist trips is not enough to comply with our information needs. Figure 4 illustrates how the concept of *HikingTrack* extends from *TouristTrip*, adding a set of additional data fields, such as total distance, slope changes, ascent, descent, aggregated climb, upper and lower height. Moreover, we introduce a set of ordered places in the track

---

[12] http://schema.org
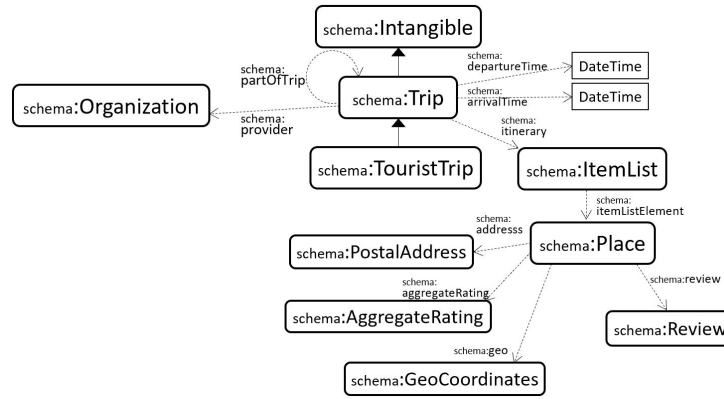
Fig. 3. Modelling a tourist trip according to schema.org. The standard trip can be used as a starting point for representing a hiking trail, as it already includes all geographical information, although it lacks specific elements such as slope information, gradient, difficulties, etc.

itinerary, which may include both interest and difficulty points. These difficulty points represent parts of the track, for which a difficulty assessment has been performed.
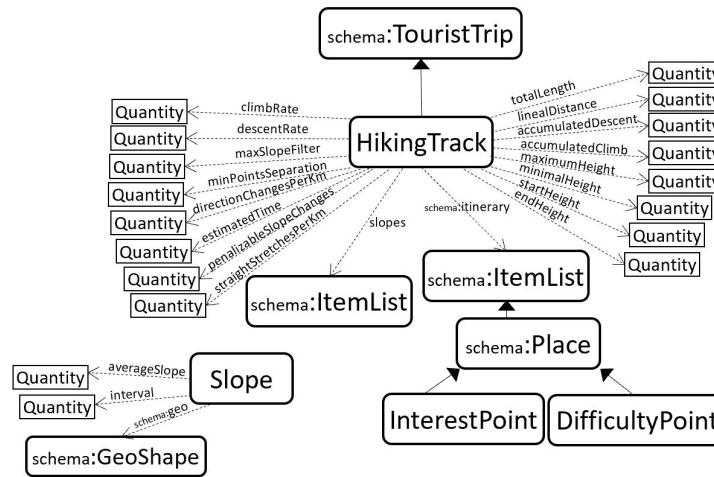
Fig. 4. Modelling of a Hiking Track, extending from schema.org. The model includes several technical information regarding the characteristics of the trail, as extracted from its geographical representation. It also includes detailed slope information and the interest and difficulty points.

The different difficulty points alongside the geographic and geometric charac-teristics of the track are essential components for the evaluation of the difficulty, as explained in the previous section. However, all this information would be in-sufficient, as they do not provide all the necessary elements with the level of granularity (i.e., difficulty types) that is needed to perform recommendations or other analytics tasks. As we can see in Figure 5, difficulty points include these additional data elements, including scores and evaluation, as well as the different aspects that we presented above: namely effort, technique, and risk.
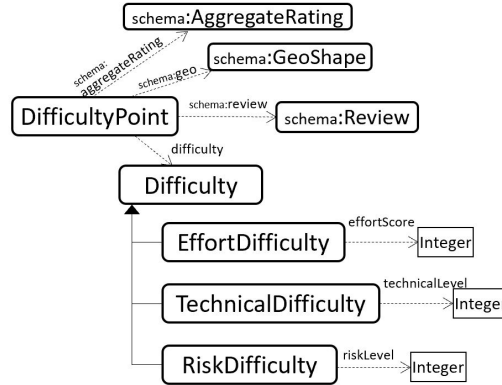
Fig. 5. Modelling Difficulty points, as extensions to a schema.org Place. the model includes the main three different types of difficulty, related to effort, risk, and technique. The model is not strongly linked to any particular type of calculation, such as the IBP index, and it could use different types of scores.

By looking at the relevant parameters on the model, the addition of these information pieces enables to answer the information questions posed above. Therefore, this allows performing comparative analysis, and potentially compar-ing track difficulties with a given set of user preferences.

The model presented in this section, implemented as an ontology in terms of Semantic Web technologies, addresses all information requirements presented in the competency questions. In particular, it allows representing hiking trails and their parts (e.g., the necessary technical information, risks, technique, and effort levels). This also reflects the different design principles detailed earlier in this paper.

6   Discussion

This paper presents an innovative conceptual model relying on knowledge-based technologies focusing on the difficulty assessment of hiking trails. Once the model

is consolidated, it will be possible to realize systems providing recommendations based on a seamless matching of the health/preferences profile of a user with the available trails. The realization of this idea lies at the intersection of disciplines that include tourism, health and well-being, and knowledge acquisition and management. Therefore, it requires a careful analysis of the needs, requirements, and perceptions of the various stakeholders involved in the subject, including regional tourist offices, local development authorities, health professionals, mountain guides, and potential hikers. The solution envisioned by this paper sets the objective of adapting the hiking offer to the physical condition and limiting factors (e.g., vertigo and lack of balance) of users. Hence, it can be an instrument of prevention for any audience: healthy people, people with reduced mobility, or people affected by chronic diseases.

The proposed model can be considered a starting point for the establishment of a trails knowledge base. To do so, each trail must be cataloged, including its different characteristics and difficulty points. The methodology to perform this task should follow the guidelines of the French Hiking Federation. To test and prepare the data acquisition phase for trail information, a series of site visits should be performed and supervised by trained guides and physiotherapist (facilitating the identification and annotation of difficulty points). For every trail, a group of people with different physical abilities and walking habits are expected to annotate the encountered difficulty points (e.g., rocks, roots on the trail, steep slopes, bridges and obstacles, vertigo points, and narrow passages). According to the model, every noted point must be associated with its geolocalization and a difficulty score.

7   Conclusions and Future work

This paper introduced a semantic model to represent hiking trail difficulties. Its goal is to create a knowledge graph for personalized recommendations and information. The presented model, based on the principles of Semantic Web modeling and ontology engineering, complies with the elicited requirements. Moreover, it reuses concepts from the well-established schema.org vocabulary, allowing potential interoperability with other third-party systems and published data sources.

Once a sufficient amount of data using this model will be collected, a hiking trail knowledge base will be released. Such a knowledge base can enable the implementation of hiking track filtering algorithms (e.g., based on machine learning techniques). Thus, a user's profile can be classified according to other users and their preferences. This automatic matching mechanism should also establish a score or scale, facilitating decision-making for the user. A feedback mechanism is also necessary to improve the recommendations based on user satisfaction and feedback. Evolving classification algorithms will be required, as well as classification techniques based on specific profile parameters. Finally, to increase the system's efficiency, the parameters' impact on the final recommendation requires further studies.

Acknowledgements

References

1. Bizer, C., Heath, T., Berners-Lee, T.: Linked data: The story so far. In: Semantic services, interoperability and web applications: emerging concepts, pp. 205–227. IGI Global (2011)
2. Boerger, M.: Hiking suggestions and planner (Apr 1 2014), uS Patent 8,688,374
3. Boller, F., Hunziker, M., Conedera, M., Elsasser, H., Krebs, P.: Fascinating remoteness: The dilemma of hiking tourism development in peripheral mountain areas. Mountain Research and Development **30**(4), 320–332 (2010)
4. Calbimonte, J.P., Zappellaz, N., Hébert, E., Simon, M., Délétroz, N., Hilfiker, R., Cotting, A.: Santour: Towards personalized recommendation of hiking trails to health profiles. In: International Conference on Web Engineering. pp. 238–250. Springer (2018)
5. Delic, A., Neidhardt, J.: A comprehensive approach to group recommendations in the travel and tourism domain. In: Adjunct publication of the 25th conference on user modeling, adaptation and personalization. pp. 11–16. ACM (2017)
6. Haskell, W.L., Lee, I.M., Pate, R.R., Powell, K.E., Blair, S.N., Franklin, B.A., Macera, C.A., Heath, G.W., Thompson, P.D., Bauman, A.: Physical activity and public health: updated recommendation for adults from the american college of sports medicine and the american heart association. Circulation **116**(9), 1081 (2007)
7. Heberlein, T.A., Fredman, P., Vuorio, T.: Current tourism patterns in the swedish mountain region. Mountain Research and Development **22**(2), 142–150 (2002)
8. Hugo, M.L.: Energy equivalent as a measure of the difficulty rating of hiking trails. Tourism Geographies **1**(3), 358–373 (1999)
9. Jackson, J., Murphy, P.: Tourism destinations as clusters: Analytical experiences from the new world. Tourism and hospitality research **4**(1), 36–52 (2002)
10. Lloyd-Jones, D., Adams, R., Brown, T., et al.: Health benefits of hiking. Circulation **121**, e1–e170 (2010)
11. Mason, R.C., Suner, S., Williams, K.A.: An analysis of hiker preparedness: a survey of hiker habits in new hampshire. Wilderness & environmental medicine **24**(3), 221–227 (2013)
12. Nguyen, T.N., Ricci, F.: A chat-based group recommender system for tourism. In: Information and Communication Technologies in Tourism 2017, pp. 17–30. Springer (2017)
13. Organization, W.T.: Thesaurus on Tourism and Leisure Activities (2001)
14. Pitman, A., Zanker, M., Gamper, J., Andritsos, P.: Individualized hiking time estimation. In: 2012 23rd International Workshop on Database and Expert Systems Applications. pp. 101–105. IEEE (2012)
15. Scarpa, R., Thiene, M., Tempesta, T.: Latent class count models of total visitation demand: days out hiking in the eastern alps. Environmental and Resource Economics **38**(4), 447–460 (2007)

16. Thiene, M., Scarpa, R.: Hiking in the alps: exploring substitution patterns of hiking destinations. Tourism economics **14**(2), 263–282 (2008)