User Engagement through Topic Modelling in Travel

Athanasios Noulas *
Booking.com
Rembrandt Square Office
Herengracht 597, 1017 CE Amsterdam
athanasios.noulas@booking.com

Mats Stafseng Einarsen †
Booking.com
Rembrandt Square Office
Herengracht 597, 1017 CE Amsterdam
mats.einarsen@booking.com

ABSTRACT

Booking.com engages its users in different ways, like for example email campaigns or on-site recommendations, in which the user receives suggestions for the destination of their next trip. This engagement is data-driven, and its parameters emerge from the corresponding relevant past behavioural pattern of users in the form of collaborative filtering or other recommender algorithms. In this work we use a secondary database with meta-information about the recommended destination in the form of user endorsements. We model the endorsements using Latent Dirichlet allocation, a well-known principled probabilistic framework, and use the resulting latent space to optimise user engagement. We demonstrate measurable benefits in two distinct interactions with the user in the form of email marketing and menu-based website browsing.

1. INTRODUCTION

Booking.com [2] is the world's largest online travel agent and millions of users visit its e-shop to reserve their accommodation. As in most shopping experiences, the users have often only partially settled their product expectations. For example, the user might have a fixed budget and dates, but still be flexible in terms of the final destination. In such cases, optimising user engagement becomes a high impact objective, as an engaged user has higher chances on settling the final details of his choice and making a purchase which leads to short- and long-term prosperity of the company.

There has been extensive work on recommending relevant items to the user [4] and much of those recommendations are provided both on the website of Booking.com as "alternative destinations" as well as through our email campaign which reaches millions of subscribers all over the world. Unfortunately, in a world-wide-web drowning in advertisements, these recommendations often fail to differentiate themselves, and they are therefore discarded. Moreover, simple collaborative filter-like messages, *e.g. user who travelled to Rome also travelled to Florence*, miss the personal touch which will engage the user optimally.

Over the last year[1], Booking.com launched the Destination Finder [1], a project that allowed users to endorse destinations for different activities. The objective of the project was to help users choose a destination based on their favourite activities rather than the location itself. What differentiates the Destination Finder endorsements from usual travel websites is that they come strictly from people who have stayed at the destination they endorse — as endorsements are associated to a completed hotel reservation. The objective of our project was to use the endorsement meta-data to optimise and personalise user engagement in every part of their interaction with Booking.com.

The remainder of this paper is organised as follows. Section 2 contains a short description of the endorsement data. Section 3 gives a brief overview of the Latent Dirichlet Allocation[3], a probabilistic topic model we used to analyse the endorsement data. Section 4 presents the results acquired on the endorsement dataset. Section 5 describes different successful applications of the resulting endorsement data model along with their quantitative comparative results.

2. ENDORSEMENT DATA

When this work was carried out the destination finder had collected more than 10 million endorsements from more than 2 million real users. These endorsements can be free, unprocessed text, or be one of 213 fixed endorsements. Some statistics of the most and least common fixed endorsements are shown in table 1.

Endorsement	# given	# destinations	# users
Shopping	513536	9105	467143
Food	242647	13595	228130
Beach	237035	8255	220603
Adventure	9622	2327	9374
Chinatown	2041	132	1992

Table 1: Example endorsements, the number of times this endorsement was given by our users, the number of destinations that have been endorsed with this endorsement and the number of users that have given this endorsement

*Data Scientist, Visitor Profiling
†Senior Product Owner, Front End

[1]The first endorsement was recorded in September 2013

KDD, The Second Workshop on User Engagement Optimization 2014 New York City, New York USA

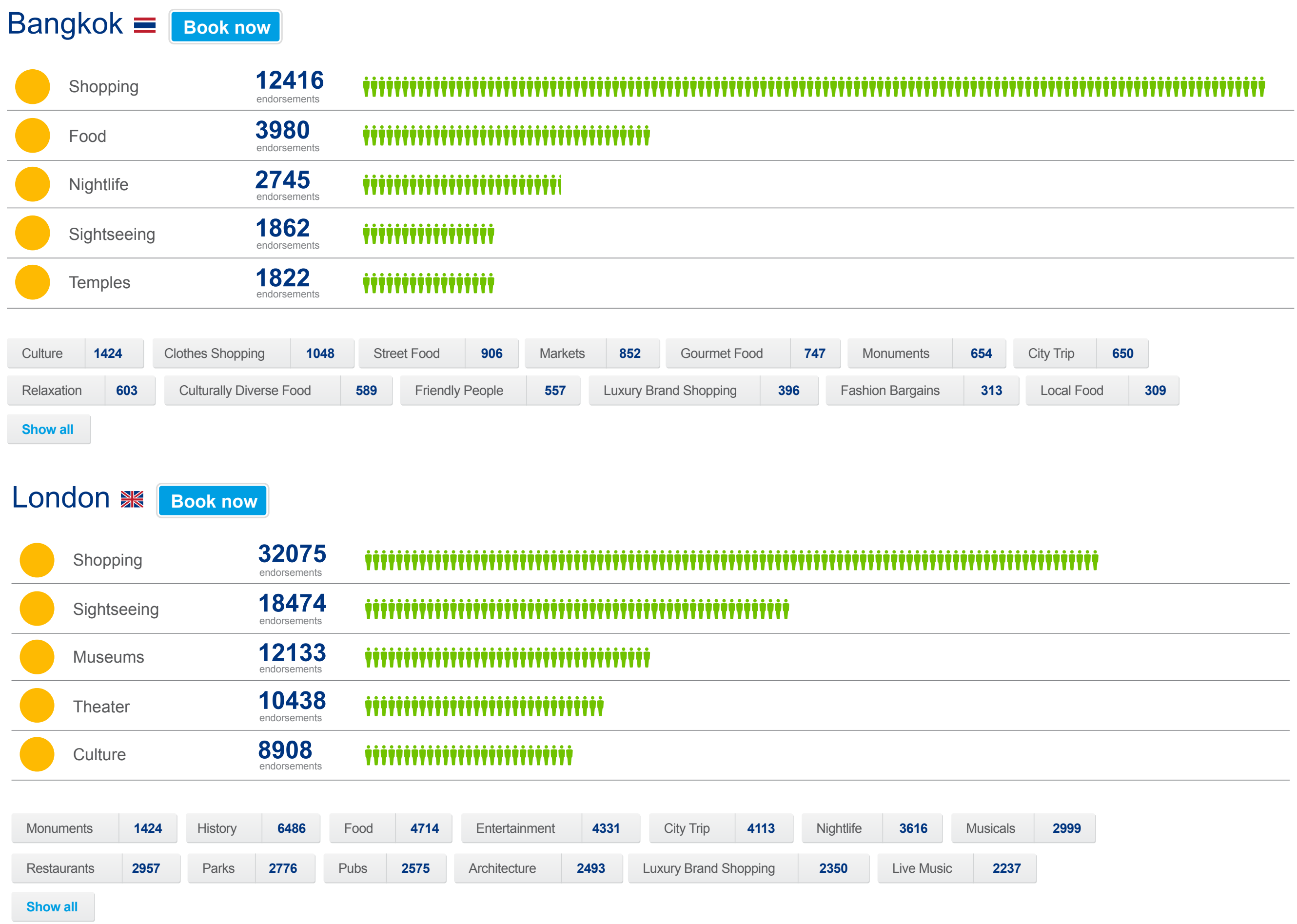

Figure 1: The endorsement pages of Destination Finder for Bangkok and London

The endorsement dataset presents a number of challenges which are common in natural language processing applications:

1. Sparsity: Most of the endorsements represent a very long tail appearing very rarely

2. Ambiguity: Endorsements like ”Shopping” can represent a variety of activities, ranging from shopping for luxury brands to buying souvenirs or bargaining in an open market

3. Competition: Most users give a limited number of endorsements, so inevitably endorsements in large cities, where multiple activities are available, will compete against each other for this limited space.

4. Redundancy: People mention ”Food”, ”Street Food” and so on interchangeably

5. Relativity: London might be a great destination for a city-trip from Amsterdam, but it is a far-away destination for someone coming from Thailand. Also, Culturally-Diverse-Food, which is a very common endorsement, has completely different meaning at different places of the world.

As a result, most cities usually have groups of endorsements with similar frequency which depend both on the city and the visitors it attracts. For example, figure 1 contains the endorsements given to Bangkok and London as they are presented in each city’s Destination Finder page. We can see that Shopping is first in both destination, while Bangkok gets higher percentage of endorsements related to nightlife in contrast to London which gets more culture-related endorsements.

3. LATENT DIRICHLET ALLOCATION

Latent Dirichlet Allocation [3] is a generative model which has been used extensively to model the distribution of words in documents. The naming comes from the fact that these topics are *latent* in the sense that we retrieve them from an unlabelled dataset as opposed to produce them using labelled data. The *Dirichlet* part refers to the priors we provide for the topic distribution in the documents and the word distribution in each topic. The objective behind the application of Latent Dirichlet Allocation is to model each document as a mixture of a small number of topics and at the same time attribute each word of the document to one of the topics present in the corresponding document.

The probabilistic graphical model of the Latent Dirich-

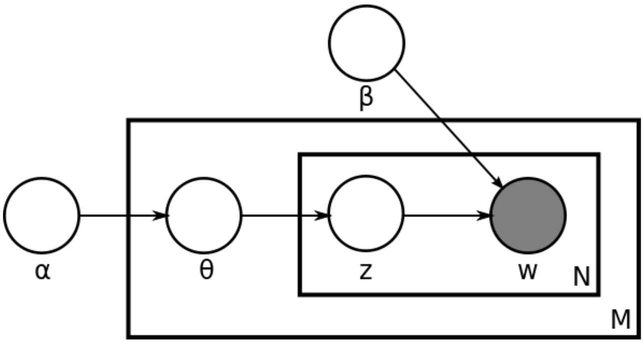

Figure 2: The probabilistic graphical model of LDA

let Allocation is visible in figure 2. The plates represent M documents each one containing some of the N words of our vocabulary. In mathematical terms α and β are priors for the Dirichlet distributions of topics and words. Each document i has a topic distribution parameterised by θ_i and each topic k has a word distribution parameterised by ϕ_k. The jth word in document i is then w_{ij} and it is associated to topic z_{ij}. The generative process then works in three steps:

1. Choose $\theta_i \sim \text{Dir}(\alpha)$, where $i \in \{1, \ldots, M\}$ and $\text{Dir}(\alpha)$ is the Dirichlet distribution for parameter α
2. Choose $\phi_k \sim \text{Dir}(\beta)$, where $k \in \{1, \ldots, K\}$
3. For each of the word positions i, j, where $j \in \{1, \ldots, N_i\}$, and $i \in \{1, \ldots, M\}$
   - Choose a topic $z_{i,j} \sim \text{Multinomial}(\theta_i)$.
   - Choose a word $w_{i,j} \sim \text{Multinomial}(\phi_{z_{i,j}})$.

In practice we need to provide the learning algorithm with a corpus containing the M documents as the counts for each of the N word for each document. We further provide the desired parameter K which represents the total number of topics we want to discover. During learning, the model will retrieve the distribution over words for each topic so as to maximise the complete corpus likelihood:

$$P(\boldsymbol{W}, \boldsymbol{Z}, \boldsymbol{\theta}, \boldsymbol{\varphi}; \alpha, \beta) = \prod_{i=1}^{K} P(\varphi_i; \beta) \prod_{j=1}^{M} P(\theta_j; \alpha) \cdot \quad (1)$$

$$\prod_{t=1}^{N} P(Z_{j,t}|\theta_j) P(W_{j,t}|\varphi_{Z_{j,t}}) \quad (2)$$

we can then use this distribution to map each of the corpus' documents to a mixture over the latent topics. We refer the reader interested in the details of learning and inference for the Latent Dirichlet Allocation model to David Blei's page [2] which contains more information regarding topic modelling, many open-source implementations of Latent Dirichlet Allocation as well as a review paper and a tutorial presented in KDD-2011.

4. MODELLING ENDORSEMENT DATA

The application of Latent Dirichlet Allocation (LDA) on our endorsement data is quite straightforward. Each endorsement is one word, while each reservation and its corresponding endorsements can be seen as a single document and its corresponding words. LDA requires a user-defined number of latent topics, and after experimentation with 5, 20, 40, 100 or 200 topics we acquired the best empirical results with 40 topics. A few sample topics are visible in table 2, along with the most common endorsements associated to each topic.

Topic	Dominant Endorsements (probability of appearance)	Example Destinations
1	Romantic (0.64), Romanruins (0.06), Photography (0.06), Historical-Landmarks (0.06)	Brugge, Vienna, Verona
2	Snorkelling (0.23), Diving (0.19), Sun-Bathing (0.11), Walking-with-Kids (0.10)	Bayahibe, Ko Phi Phi Don, Rhodes
4	Shopping (0.40), Food (0.35), Walking (0.07), Entertainment (0.06)	Taipei, Buenos Aires, Valencia
23	Shopping (0.28), Monuments (0.12), Culturally-Diverse-Food (0.10), Business (0.08)	Sydney, Lisbon, Liverpool

Table 2: Some of the topics discovered by the Latent Dirichlet Allocation. The topic number an identifier of the latent topics discovered by LDA. The dominant endorsements are the endorsements that appear most probably when a trip is associated to the corresponding topic, while in brackets we have their probability of appearance.

At this moment there is no topic modelling algorithm universally accepted as the optimal solution. We found LDA to be very efficient for our endorsement data, but as with any application on a real-world dataset there were some interesting observations which we organise here in three sections, namely the good, the bad and the ugly.

4.1 The good

The LDA allows us to map any trip to a mixture of topics, since during training we used each individual trip as a document. Moreover, we can express each destination as a mixture of the trips done there and thus map each destination to the latent topic space.

In the example of Bangkok and London the main topics are visible in table 3. In both destinations, as in any big city, Shopping is one of the most common activities travellers do, but in this case it is disambiguated into different topics. People who travel to London combine shopping with Sightseeing and Theater, or do it as part of the Christmas Shopping, while people in Bangkok combine it with Food and Entertainment or monuments and Culturally-Diverse-Food. Finally, the two destinations are clearly separated by their second most prominent topic which in case of London is Museums, while in the case of Bangkok it is Culture and Temples.

Mapping destinations to topics works extremely well for major destinations since we have thousands of trips to describe them. This mapping also works very well for less popular destinations, discovering great places for niche endorsements like "Fine-Dining", "Volcanoes" or "Bar-hopping". Moreover, it is very easy to express popularity by multiplying the

[2]http://www.cs.princeton.edu/ blei/topicmodeling.html

City	London			Bangkok		
Topic	25 (0.29)	10 (0.10)	18 (0.097)	4(0.27)	35 (0.13)	23 (0.11)
End.	Shopping (0.40)	Museums (0.8)	Shopping (0.78)	Shopping(0.40)	Culture(0.30)	Shopping(0.28)
End.	Sightseeing (0.36)	Parks (0.03)	Christmas (0.05)	Food (0.35)	Temples (0.24)	Monuments(0.12)
End.	Theatre (0.36)	Galleries (0.03)	Culture (0.04)	Entertainment(0.06)	Food (0.20)	Diverse Food (0.10)

Table 3: Main topics and their endorsements for London and Bangkok

topic distribution of a destination with its total number of visitors, or discover trends by finding relative changes in the topics of each month's endorsements.

4.2 The bad

Similarly we can map users to topics based on the endorsements they provided for their past trips. Unfortunately, the travel industry has notoriously sparse data. Netflix users might watch tens or hundreds of items in a year, while our average user performs just below two trips. Moreover, users often do not give endorsements feedback on their trips.

We, therefore, model each user's preferences as the product of the topic distributions of the destinations they visit. A different choice would be modelling the user as a mixture of their visited destinations, but we are interested to the single topic that will engage them the most, rather than all the topics they have been exposed to.

On the destination dataset, there exists a set of destinations that have a very small number of endorsements with respect to the total number of visitors they get — like for example airports and other travelling hubs. The small number of endorsements can produce a very peaked distribution in the topic space, while the high number of visitors might bring destinations like Heathrow or Schiphol on the top of the respective topic's lists. We address this issue by blacklisting the 1‰ of the destinations with the worst endorsements to visitors ratio.

4.3 The ugly

Endorsement data are user generated and so are the fixed endorsements we used for the LDA model. Inevitably, there exist terms which might not be socially acceptable by some cultures or can be found offensive in others, like for example the endorsement *People Watching* which was given approximately 6000 times. As long as we stay within the most common endorsements, we are relatively safe, but using the full width of our vocabulary requires careful inspection of user engagement.

The idea behind Latent Dirichlet Allocation is to map a large number of words to a small number of topics. This means that many words will be grouped together, and in some topics, these activities might be very diverse. For example Topic 17 contains Culturally-Diverse-Food (0.43), Theatre (0.35), Modern-Art (0.09) and Scuba-Diving (0.05). We need to be careful before using these endorsements simultaneously, as the final result might be confusing to the user.

Last but not least, LDA returns a number of topics with no clear naming. The success of our experiments will depend on the name we give to the topics ourselves, and simple solutions like using the top endorsement for each topic might take away the content discovered by the topic modelling — see for example different Shopping topics in table 3.

Eng.	Users	Interaction	Net Conversion	Cancellations
Eml 1	34M+	**+18.34%**	**+10.00%**	**+14.80%**
Eml 2	34M+	**+18.71%**	**+7.14%**	**+4.39%**
Menu	40M+	Ongoing Experiments		

Table 4: Quantitative Results of the user engagement campaigns. Users as split randomly between the base and variant groups, and bold indicates statistically significant results.

5. USER ENGAGEMENT APPLICATIONS

The probabilistic nature of the resulting endorsement model makes many different applications feasible. We experimented in multiple areas of interaction with the user to optimise their engagement to our product and facilitate their accommodation search, *e.g.* email marketing, product organisation and recommendation personalisation. All our experiments were A/B tested for statistical significance, and in almost all of our experiments we saw a positive response to our KPIs.

In this section we present three experiments in which we observed statistical significance in net conversion in. The experiments' results in two main KPIs, namely *Net Conversion* and *Cancellations*, are listed in table 4. The increase in net conversion is

$$\text{Net Conversion} = \frac{\text{Net Conversion in B}}{\text{Net Conversion in A}} - 1 \qquad (3)$$

where *Net Conversion* is the probability a user engaged will reserve an accommodation in our website and stay at the reserved accommodation.

The *Cancellations* KPI is the probability that a customer who reserved an accommodation will later cancel it — and we try to keep this ratio as small as possible — although a more engaging campaign produces a higher cancellers ratio.

5.1 Email Marketing

Email campaigns are a direct and economical way to expose our users to our products. In the case of booking.com, these emails contain proposed travelling destinations and information about special deals from partner hotels. Both personalised recommendations and deals have an obvious value to the user, however, the repeating nature of the email makes the campaigns wear-off over time.

5.2 Email 1

We used the latent topics discovered through the Latent Dirichlet Allocation analysis to detect the common topics between the destinations visited by the target user and the destinations returned by our recommender. A sample of the resulting email is visible in figure 3, in which case the main endorsements of the topic selected for the target user are Beach, Relaxation and Food.

We performed A/B testing on more than 34 million users, half of whom received the most popular endorsements email

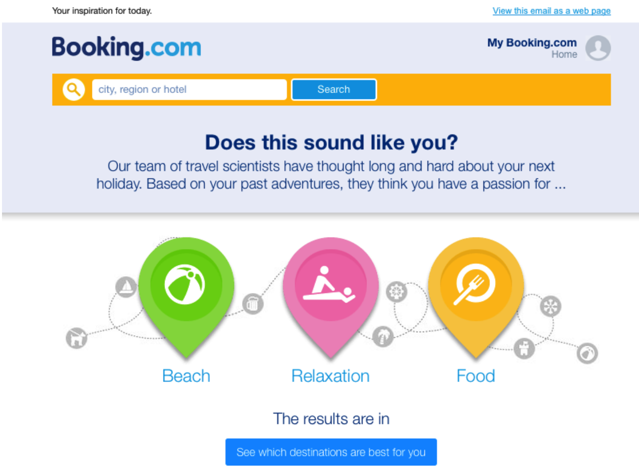

Figure 3: The LDA-based email

(group A) and half the email with their personalised LDA-based topic selection (group B).

The campaign was very beneficial in multiple KPIs, including users clicking on the campaign, interacting with the website and finally converting from visitors to bookers. More specifically, We observed a 18% uplift in clicks at the same bounce rate, 10% uplift in net conversion and 14% increase in cancellations (22% increase in gross conversion) . All of these differences were statistically significant for a confidence interval of 90%.

5.3 Email 2

The first email campaign used a quite bold statement in the form of "our team of travel scientists ... based on your past adventures they think that you have a passion for... ", visible in figure 3. This spawned a vivid response from the community on twitter with people being excited about our campaign, people complaining about the recommendations they got and people being sceptical about the access we have to their past data 4.

Our original message could have been misleading, as we didn't look at a person's activities or endorsements but rather the most common activities in the places they had visited, as explained in section 4.2. We launched a second email campaign where we used a less intrusive message promoting endorsements in suggested destinations without a direct reference to how these endorsements were selected.

We sent the email to more than 22 Million users with half of them (group A) receiving the old text and half of them (group B) receiving the new text. We saw an impressive increase in net and gross conversion (+7.14% and 10.5%) and a modest increase in cancellations (+4.39%).

5.4 Browsing Menu

The Destination Finder provided our users with a search engine that allowed them to find the ideal destination for their desired activity. Although this method offers complete freedom of search to the user, it requires them having a concrete activity in mind before coming to the website. In contrast, most e-shops organise their inventory in intuitive hierarchies, and users are familiar with menu-based navigation.

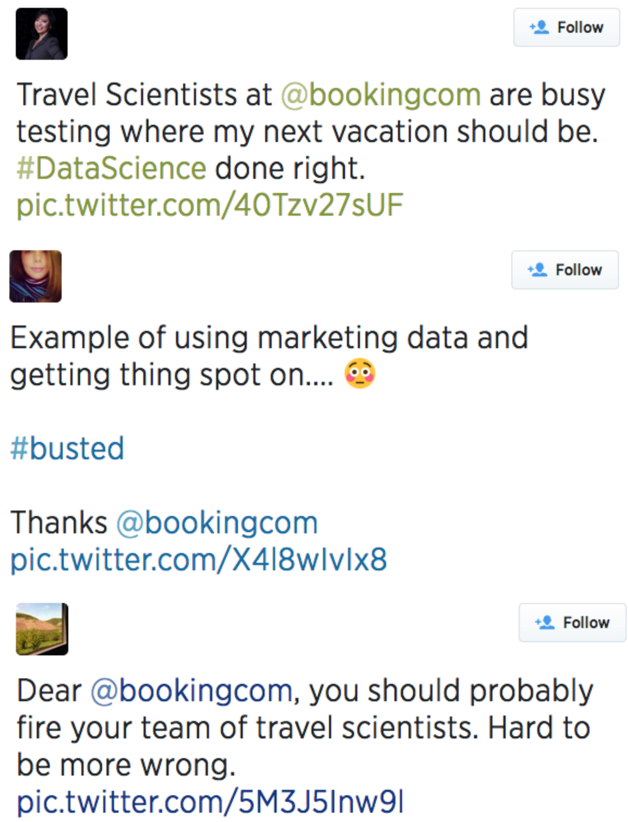

Figure 4: Example of twits related to the latent topic modelling campaign.

We produced country-specific menus, where we present as a top menu hierarchy the topics of destinations which bookers from the corresponding country prefer the most. For example, people from the Netherlands get top-level menus for Shopping, Fine Dining, Cycling, Sightseeing and Beach, as opposed to Beach, Food, Historical Landmarks, Museums and Shopping for Britain.

The second layer provides the countries where these activities are carried out most often. For example, people from the Netherlands and Britain share destinations like Spain, Italy and Greece for the topic Beach, but they only receive their individual home-country in the second level of their menus. Lastly, each of the countries opens a third level with specific destinations in this country. An example of the menu for a user from the Netherlands is visible in figure 5

The menu-based navigation was a major success. We exposed more than 40 million users in a A/B test experiment and observed statistical significance in all the user engagements KPIs we had set. Unfortunately, since the menu improvement is an ongoing project at our website at this moment, we can not publish numerical results.

6. CONCLUSIONS AND FUTURE WORK

We have demonstrated a successful application of topic modelling of user generated endorsement data using Latent Dirichlet Allocation. The automatically discovered topics have semantic significance and they can be used to optimise the user engagement experience, for example in the form

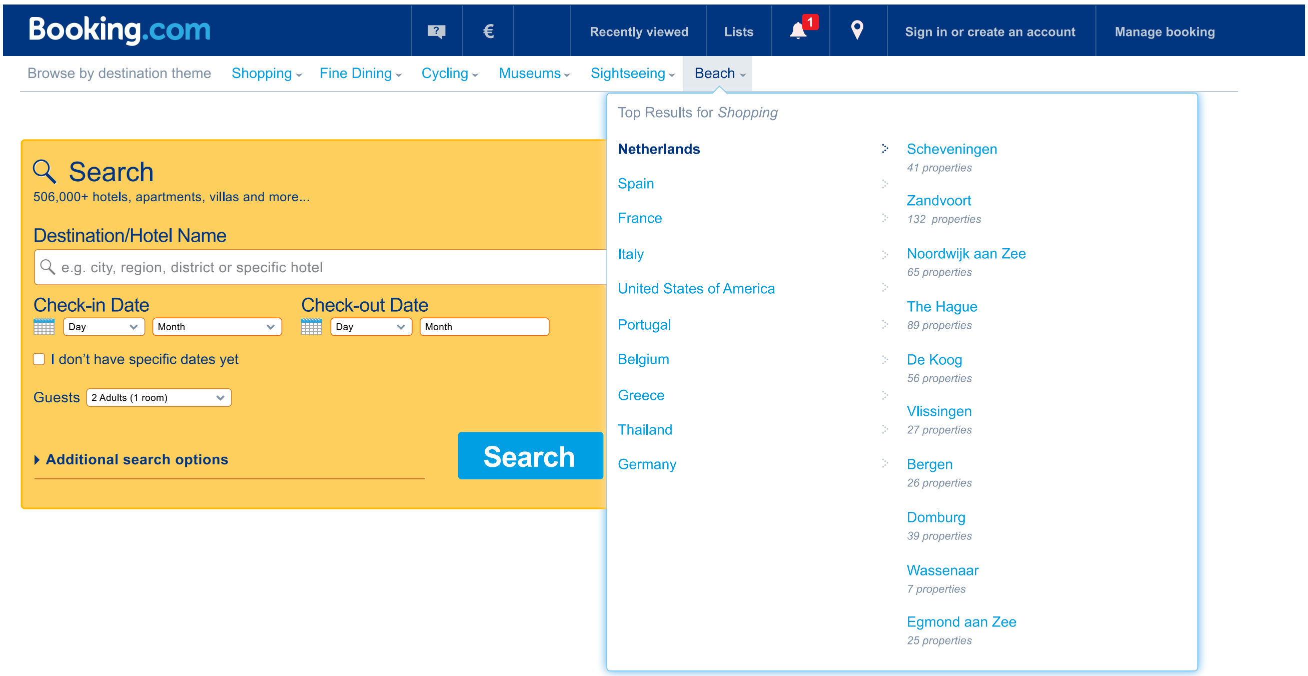

Figure 5: The LDA-based menu

of personalised email marketing and menu-based product browsing.

The initial results are very promising but we are still far away from exploiting the potential of the latent topic space. Our destination recommendations are still based on algorithms that ignore the underlying topic distribution, and the effect that the individual topics (and their manually-selected names) have on the users is hardly accounted for.

At the moment we monitor the individual interactions of each latent topic to each cluster of users in our database, in order to use them not only according to their likelihood but also according to their effectiveness. Moreover, we work on an email campaign based on consecutive follow-ups dependent on the user's responsiveness, *e.g.* using the second best topic choice if the first one failed to engage the user.

7. ACKNOWLEDGEMENTS

We would like to thank Stefano Baccianella, Gabriel Agu for their work on the data processing for the Latent Dirichlet Allocation and production implementation of the menu-based navigation software, as well as Rozina Szogyenyi and Leandro Pinto for their work on the Destination Finder and Fabrizio Salzano for overseeing the implementation of the email marketing campaign.

8. REFERENCES

[1] Destination finder webpage. `http://www.booking.com/destinationfinder.en-gb.html`. Accessed: 2014-07-10.

[2] Booking.com webpage. `http://www.booking.com/destinationfinder.en-gb.html`. Accessed: 2014-07-10.

[3] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3:993–1022, Mar. 2003.

[4] F. Ricci, L. Rokach, and B. Shapira. Introduction to recommender systems handbook. In *Recommender Systems Handbook*, pages 1–35. 2011.