Hotels.com's journey to becoming an Algorithmic Business

Matthew Fryer

VP, Chief Data Science Officer

mfryer@hotels.com

Hotels.com

Hotels.com

Part of Expedia, Inc. family

385,000 properties

89 countries

39 languages







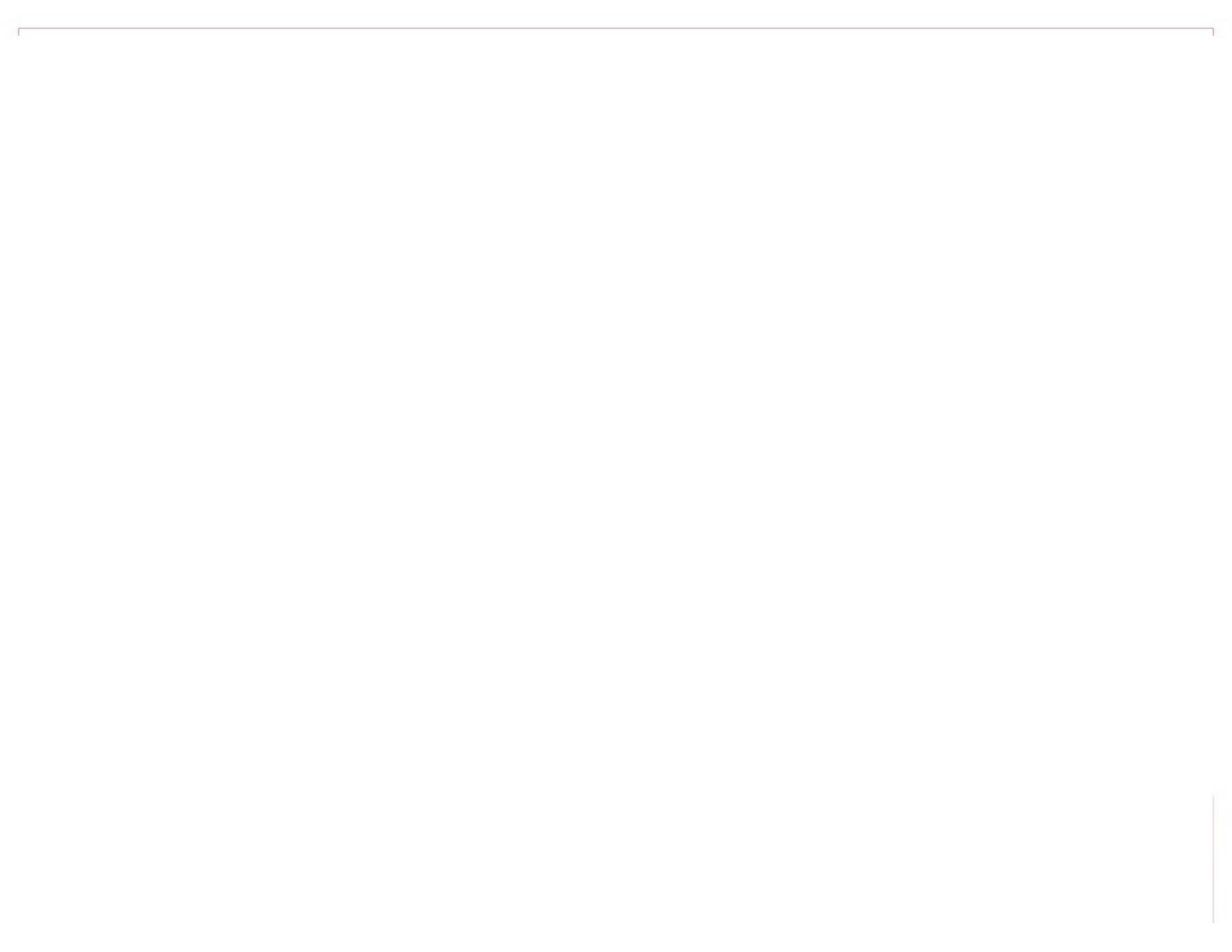


>27m Hotels.com Rewards Members

Home of Captain Obvious

Billions of Recommendations, based on real-time Data per day











Data Science



Engineering



Front End Development







Barry Diller Chairman & Senior Executive, Expedia, Inc. "Artificial Intelligence Will Be Travel's Next Big Thing"

3M's are disruptive technology

Mobile

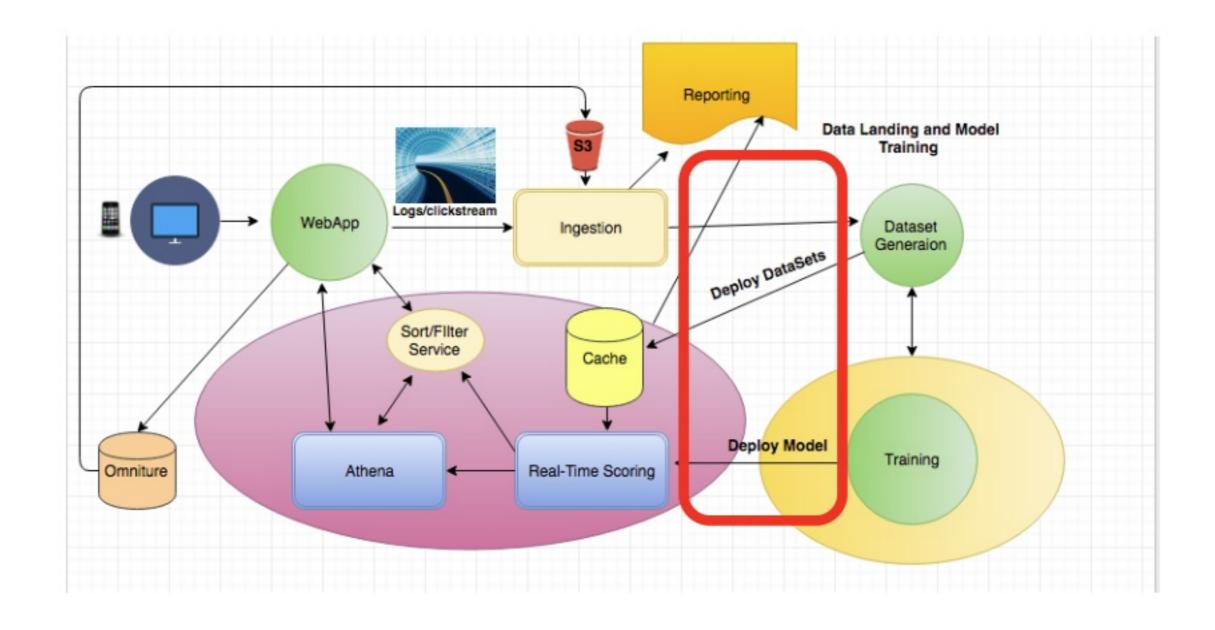
Messaging / NLP

Machine Learning





Our overall ecosystem

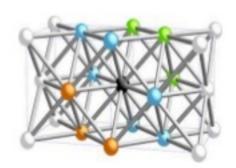


Core Elements of our Data Science Cloud Platform

Databricks Unified Platform

Maestro – Our Internally Developed Platform on AWS

(EMR, Spark, R-Studio, Intellij, SBT, Jupyter, Zeppelin, Unit / QA, Metastore, Apache Airflow, Keras, Tensorflow) Proof of Concept on Google Cloud, Beam, Spark & Tensorflow





Learning Engine



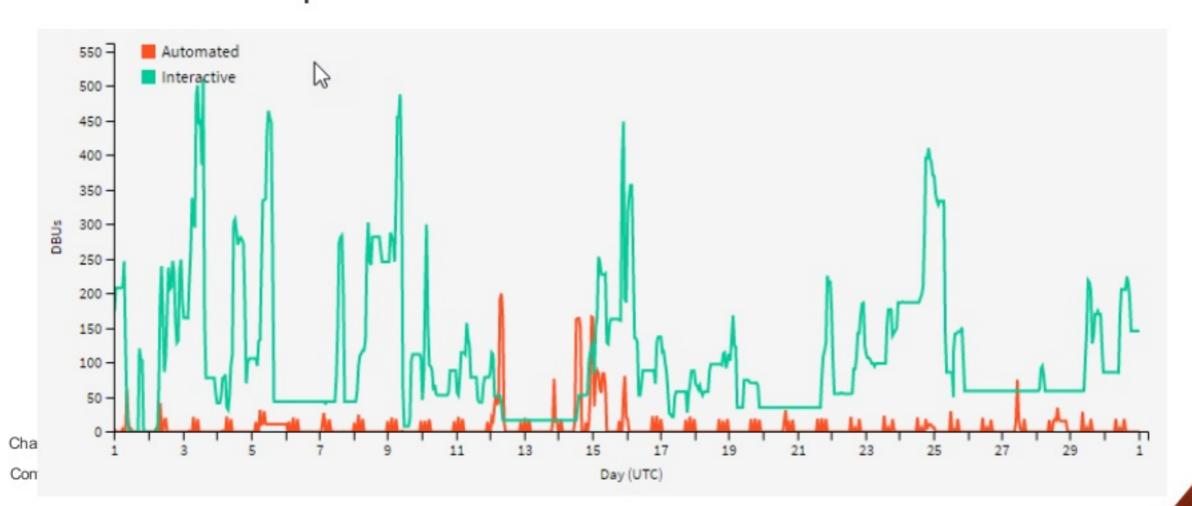






Databricks Unified Platform

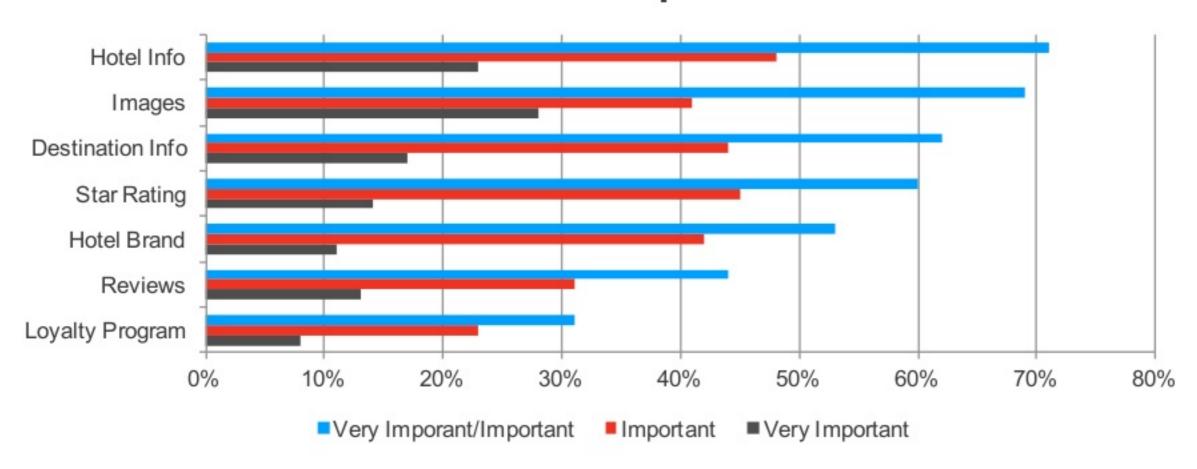
- Key asset to the success of data science at Hotels.com
- Key in driving up data scientist productivity / efficiency / flexibility
- Helps make our data science lifecycle operate much easier and faster driving speed to market
- Reliable / secure + facilitates 'Highly Elastic' workflows exploiting cost effective spot instance on AWS.





Images are an important factor while choosing a hotel

Factors other than price/location



Reference: The Influence of Visuals in Online Hotel Research and Booking Behaviour

Computer Vision problems we try to tackle



Near Duplicate Detection



Scene Classification

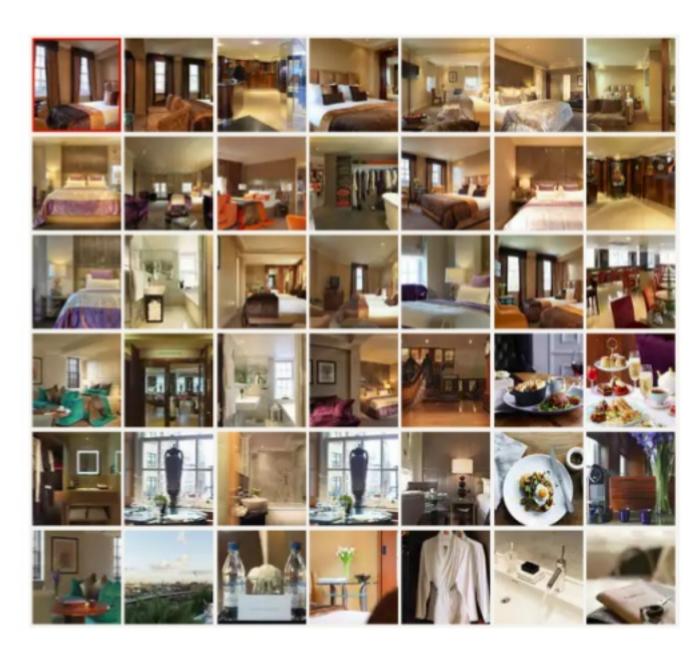
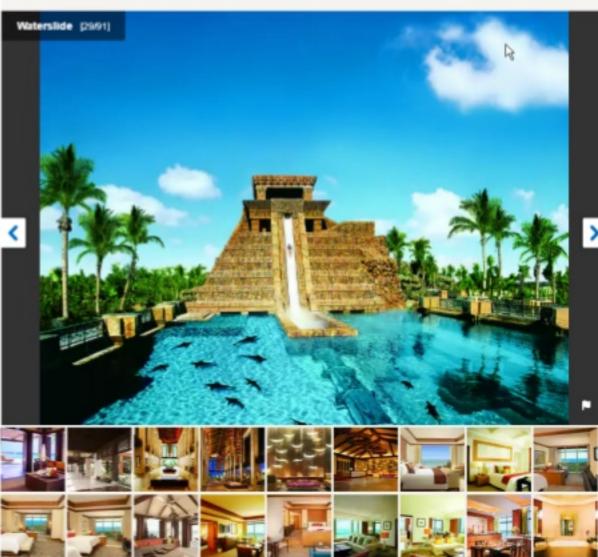


Image Ranking







Lazise 0.51 mi from Bardolino 020 3564 4852 (Call cost)

Superb! 4.5/5 (6 reviews) 2 nights from 1 left at £201



Hotel Remàt ★★★
Garda (Garda)
0.64 mi from Bardolino
020 3564 4852 (Call cost)
✓ Free Cancellation

Superb! 4.5/5 (51 reviews) 2 nights from 3 left at £202

+2.00 EUR due at hotel

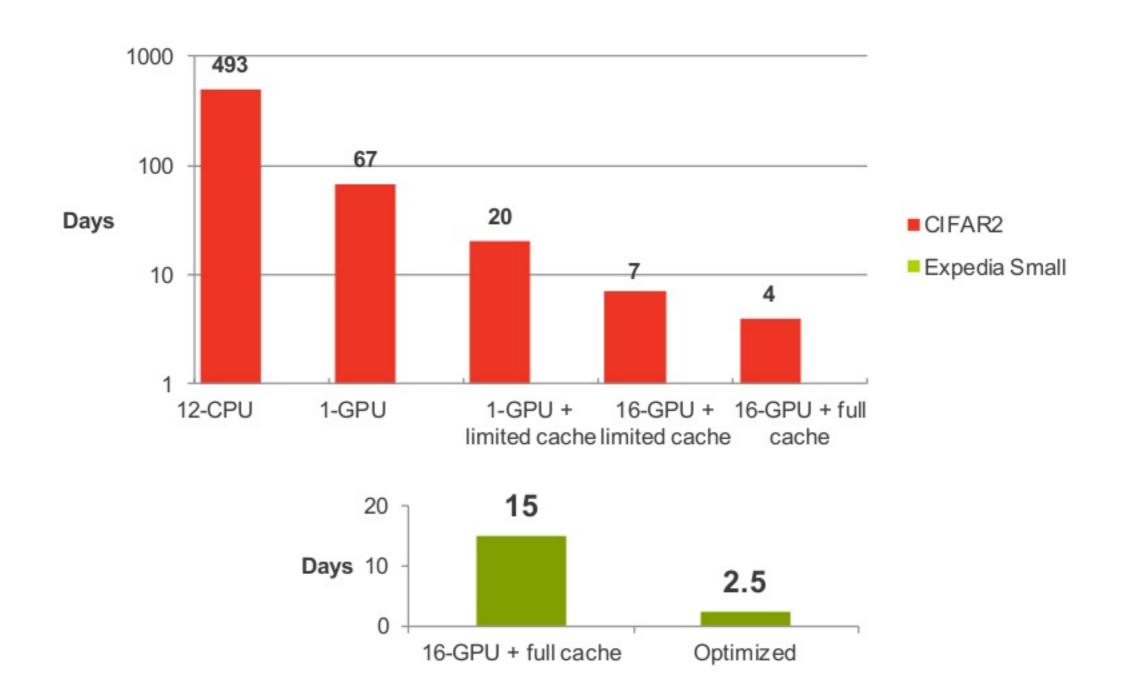


Tagged as Bathroom

Confidential - do not distribute

Hotels.com

GPU's quickly became key, took a large effort to optimize using Keras + Tensorflow (Inception v3 + ResNet)



Near Duplicate Detection: Real world examples





Non-Duplicates – probability 100%



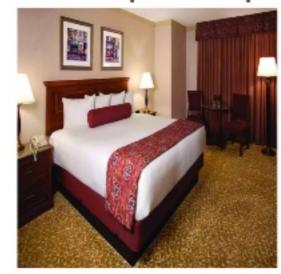


Non-Duplicates - probability 95.91%





Duplicates - probability 97.98%





16

Duplicates – probability 98.43%

Using the model: Real world examples



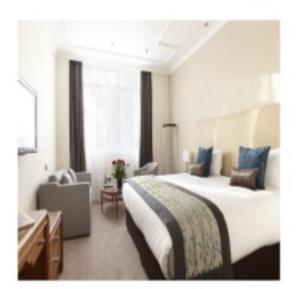
EXTERIOR/HOTEL



INTERIOR/SEATING_LO



INTERIOR/SEATING_LOBBY



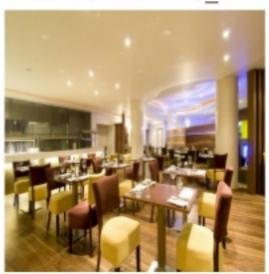
ROOM/GUESTROOM



ROOM/BATHROOM



ROOM/LIVING_ROOM

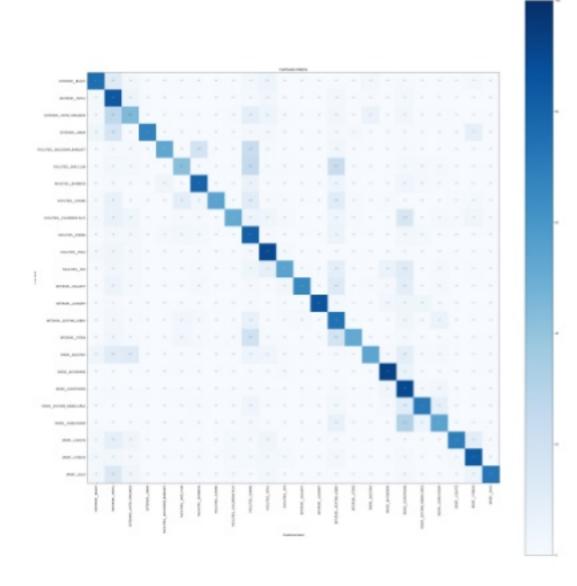


FACILITIES/DINING



FACILITIES/POOL

Accuracy & Confusion Matrix



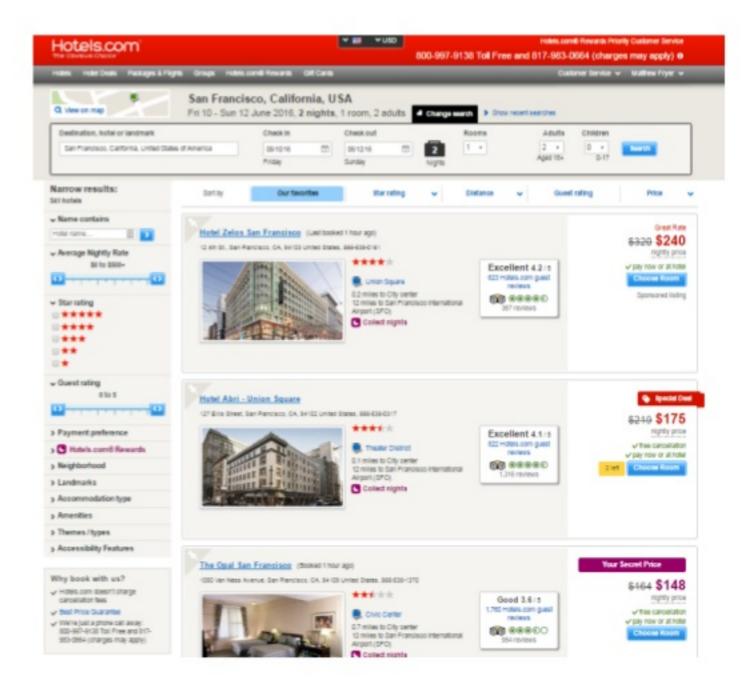
- After many manual / long winded iterations and regularization processes tuning hyperparameters
- We achieved good accuracy and low confusion matrix

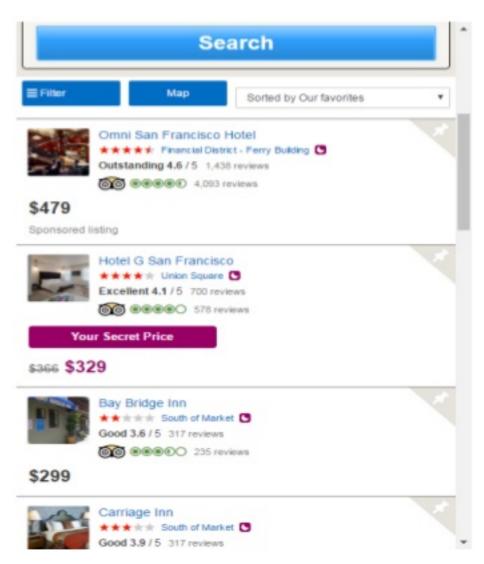
Optimizing the photo order for improved customer experiences



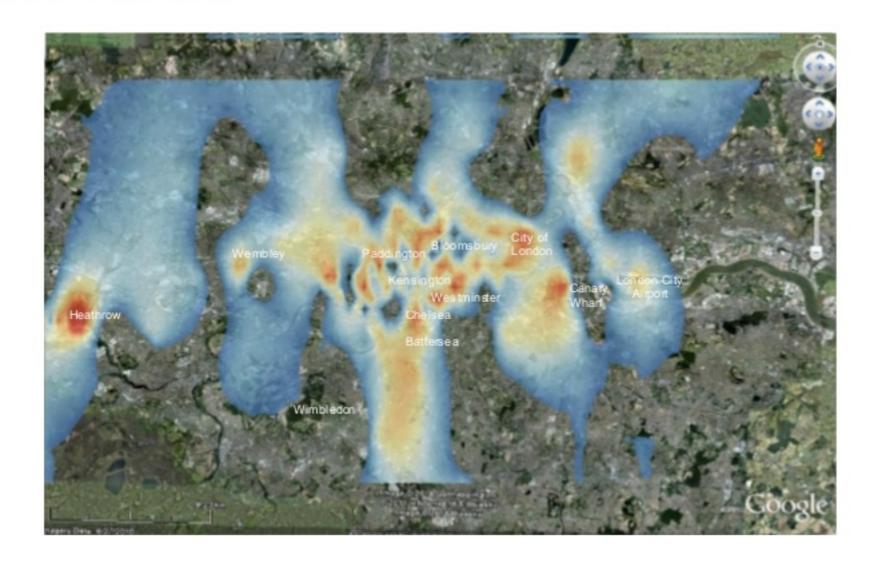
Reference: Radisson Blu Edwardian Berkshire Hotel, London

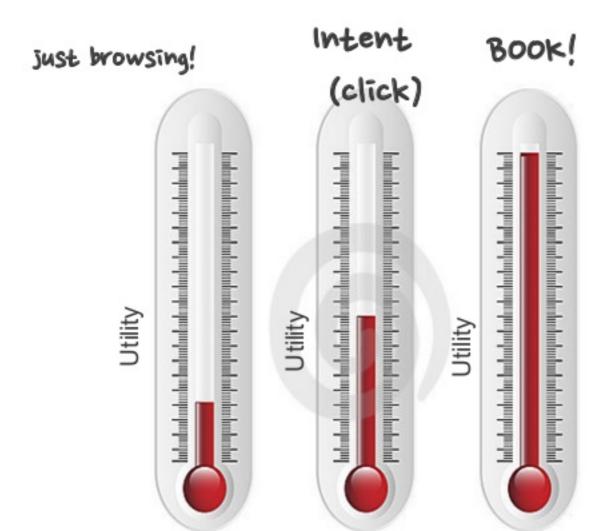
Finding the right hotel in our marketplace is core to our customers needs.





As an example different user segments like to stay in different locations







22

Thank you

mfryer@hotels.com

https://uk.linkedin.com/in/matthewfryer

@mattfryer

Hotels.com