Booking.com

Machine Translation at Booking.com Journey and Lessons Learned

May 30, 2017, Prague

Pavel Levin
Nishikant Dhanuka
Maxim Khalilov

EUROPEAN
ASSOCIATION
for MACHINE
TRANSLATION

Who am I?

About me

- Master in Computer Science (NLP) from IIT Mumbai
- 8 years of work experience in analytics and consulting
- **Data Scientist at Booking.com** since last 2 years
  - Partner Services Department (Scaled Content)
- linkedin.com/in/nishikantdhanuka/

About Booking.com

World's #1 website for booking hotels and other accommodations
- Founded in 1996 in Amsterdam; part of The Priceline Group since 2005
- 1,200,000+ properties in more than 220+ countries; 25 Million rooms
- Over 1,400,000 room nights are reserved on Booking.com every 24 hours
- Employing more than 15,000 people in 199 offices worldwide
- Website available in 43 languages

Booking.com

Agenda.

- **Motivation**
  - ☐ MT critical for Booking.com's localization process

- **MT Journey and Lessons Learned**
  - ☐ MT Model & Experiments
  - ☐ Evaluation Results
    - ◦ Automatic & Human
    - ◦ Sentence Length Analysis
    - ◦ A/B Tests
  - ☐ Interesting Examples

- **Conclusion and Future Work**
- **Q & A**

Booking.com

Motivation

Mission: Empower people to book any hotel in the world, while browsing high quality content in their own language.

2/3 of daily bookings on Booking.com is made in a language other than English

... thus it is important to have **locally relevant content at scale**

How Locally Relevant?

- Present **Hotel descriptions** in the language of the user
- Allow partners and guests to **consume and produce content in their own language**
  - ⍰ Customer Reviews
  - ⍰ Customer Service Support

Why At Scale?

- **One Million+ properties** and growing very fast
- **Frequent change requests** to update the content
- **43 languages** and more
- New **customer reviews / tickets** every second

Booking.com

Currently **Hotel descriptions** translated by human **in 43 languages** based on **visitor demand.**
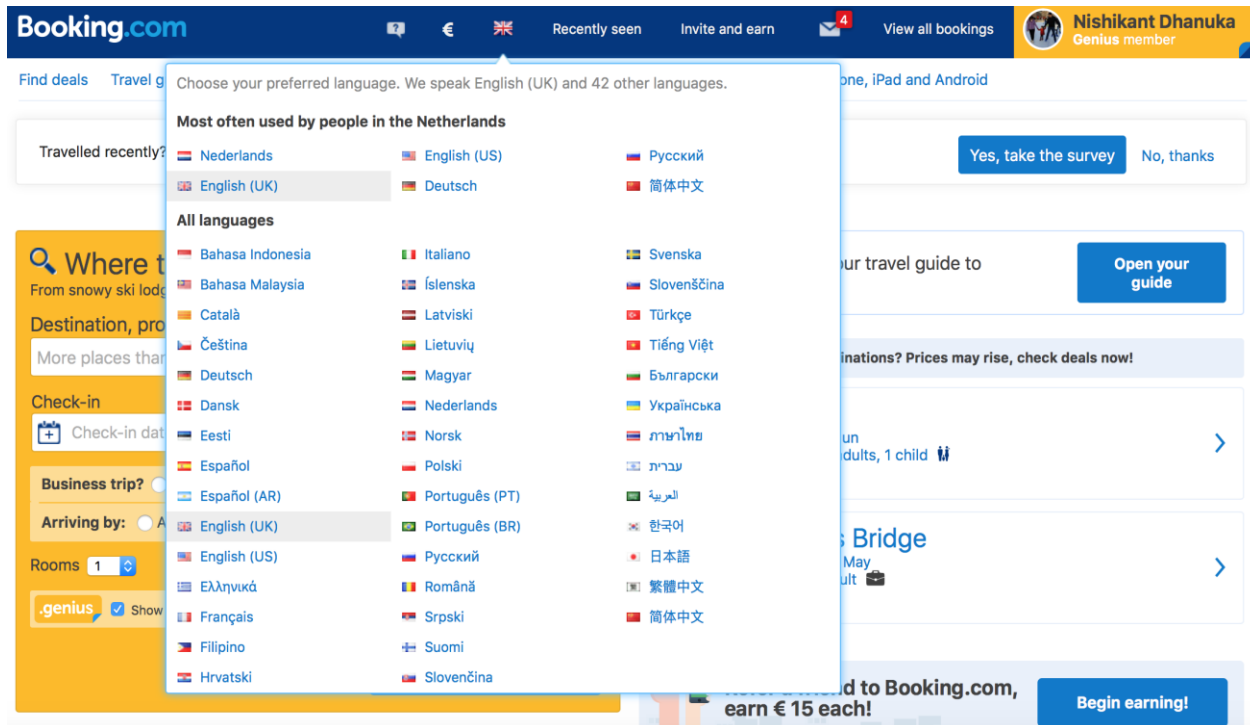
*50%

Translation Coverage

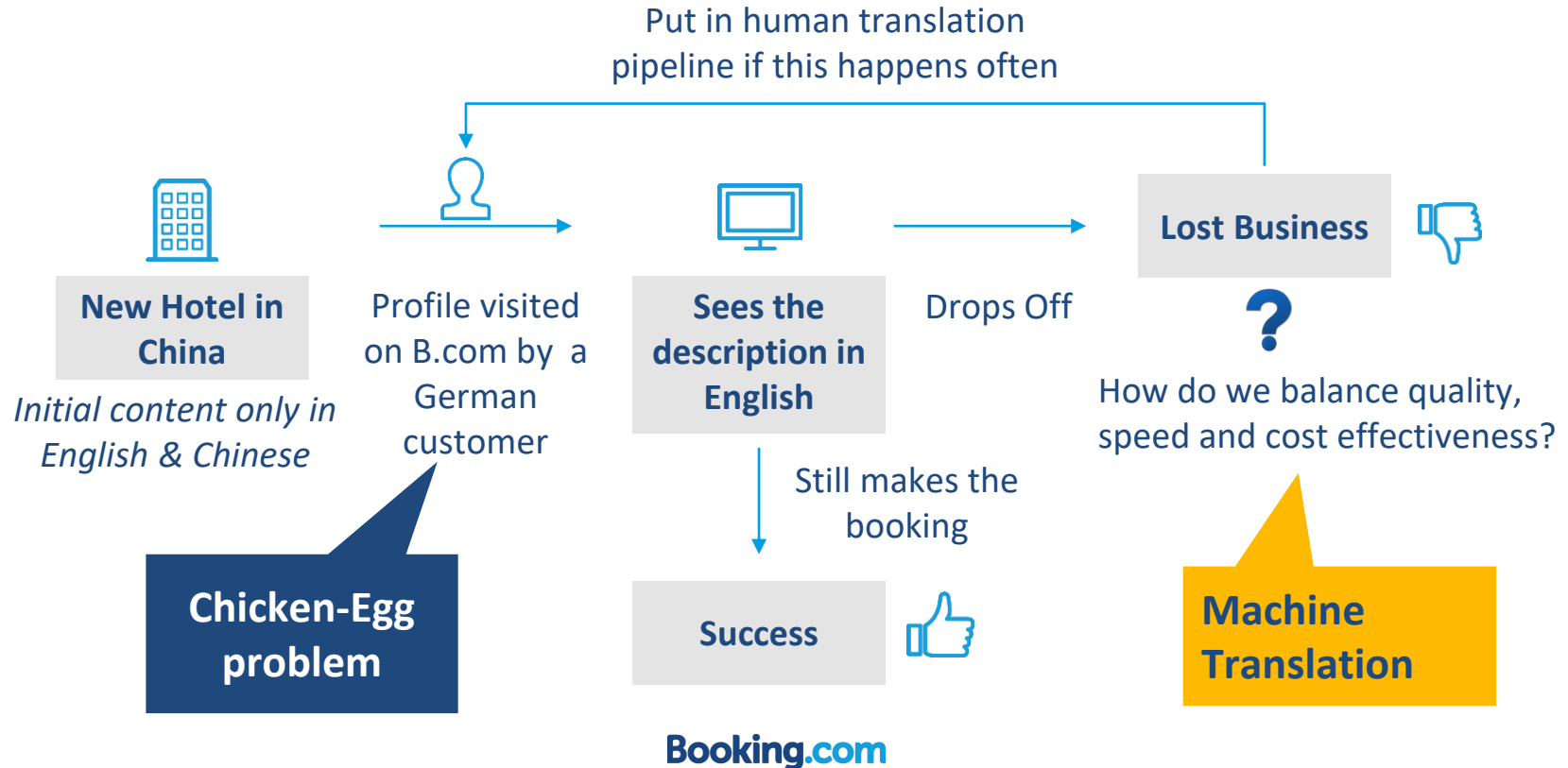***90%**

Demand Coverage

* Approximate numbers based on average of some languages

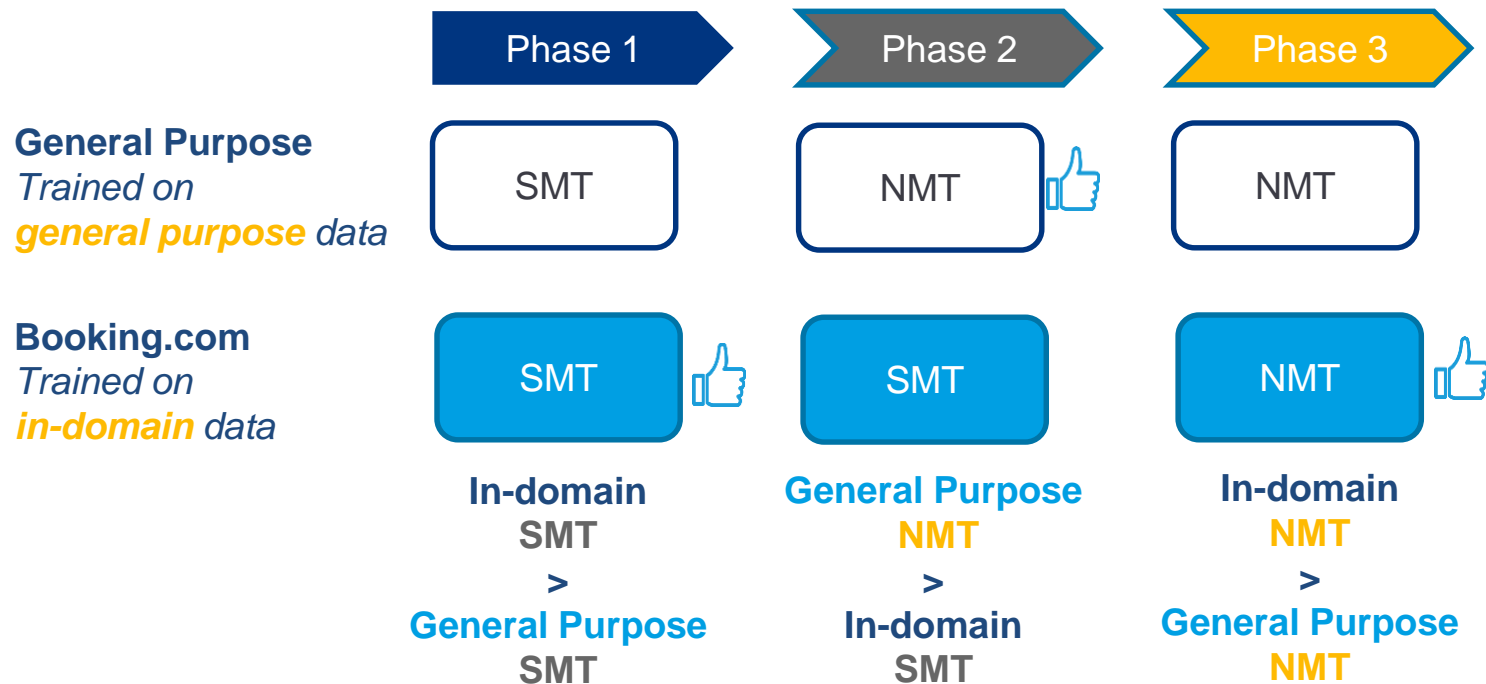# Example of Lost Business Opportunity because of highly manual and slow process.

Put in human translation
pipeline if this happens often

New Hotel in China

Initial content only in English & Chinese

Profile visited on B.com by a German customer

Sees the description in English

Drops Off

Lost Business

?

How do we balance quality, speed and cost effectiveness?

Still makes the booking

Success

Chicken-Egg problem

Machine Translation

Booking.com

MT Journey and Lessons Learned

Our Journey to discover the awesomeness of NMT

	Phase 1	Phase 2	Phase 3
General Purpose *Trained on* **general purpose** *data*	SMT	NMT 👍	NMT
Booking.com *Trained on* **in-domain** *data*	SMT 👍	SMT	NMT 👍

In-domain
SMT
>
General Purpose
SMT

General Purpose
NMT
>
In-domain
SMT

In-domain
NMT
>
General Purpose
NMT

Booking.com

Lots of in-domain data to train the MT system

Age of child on 4 Jun 2017
2
.genius ☑ Show Genius discounts first
Search

OLD TOWN
Lennon Wed PRAGUE 1
Show on map
Map data ©2017 Google

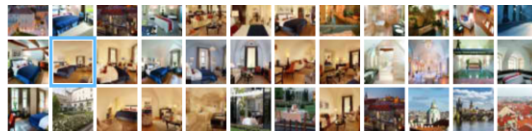How to get to Mandarin Oriental, Prague from Vaclav Havel Airport Prague

🚌 Public shuttle — 50 minutes
🚌 Private shuttle — 40 minutes
🅿 Parking available

Like this one, but not quite sure yet?
Show similar properties

Stay in the heart of Prague – Excellent central location – show map

⭐ One of our top picks in Prague.

Occupying a carefully restored monastery dating back to the 14th century, the Mandarin Oriental, Prague is located in the historic Mala Strana quarter below Prague Castle. Charles Bridge is a 5-minute walk away.

The rooms at the Prague Mandarin Oriental feature underfloor heating, and guests can choose from various bed linen and pillows. Maid service is offered twice daily, including an evening turn-down. Bathroom cosmetics include carefully selected essential oils.

The unique spa is the only one in the world located in a former Renaissance chapel. It offers a wide range of spa treatments, yoga classes and a beauty room.

Spices Restaurant and Bar invites guests to a world of authentic Asian flavours and aromas. Asian fares along with original cocktails, wine and tea can be served either at the dining table in the restaurant, at the bar or in a private booth. Spices Lounge and a covered outdoor terrace offer light relaxation and refreshment such as salads, traditional Asian bento boxes as well as classic beef burgers.

The Hellichova Tram Stop is 100 yards away from the Mandarin Prague.

Language	Parallel Sentences	# of Words	Vocab Size	Avg. Len
English -> German				
German	10.5 M	171 M	845 K	16.3
English		174 M	583 K	16.5
English -> French				
French	11.3 M	193 M	588 K	17.7
English		188 M	581 K	16.7

Booking.com

Our NMT Model Configuration Details

Data Preparation

Split Data	Train, Val, Test
Input Text Unit	Word Level
Tokenization	Aggressive
Max Sentence Length	50
Vocabulary Size	50,000

Model

Model Type	seq2seq
Input Embedding Dimension	1,000
RNN Type	LSTM
# of hidden layers	4
Hidden Layer Dimension	1,000
Attention Mechanism	Global Attention

** Approx. 220 Million Parameters

Training

Optimization Method	Stochastic Gradient Descent
Initial Learning Rate	1
Decay Rate	0.5
Decay Strategy	Decrease in Validation Perplexity <=0
Number of Epochs	5 - 13
Stopping Criteria	BLEU + sensitive sentences +constraints
Dropout	0.3
Batch Size	250

** 1 Epoch takes approx. 2 days on a single NVIDIA Tesla K80 GPU

Translate

Beam Size	30
Unknown Words Handling	Source with Highest Attention

Evaluate

Auto	BLEU WER
Human	A/F
Other	Length A/B Test

** MT pipeline based on Harvard implementation of OpenNMT

1. Data Preparation: Tokenization and Vocabulary

Data Preparation	
Split Data	Train, Val, Test
Input Text Unit	Word Level
Tokenization	Aggressive
Max Sentence Length	50
Vocabulary Size	50,000

Aggressive only keeps sequences of letters / numbers i.e. doesn't allow mix of alphanumeric as in: "E65", "soft-landing"

EN: The rooms at the Prague Mandarin Oriental feature underfloor heating, and guests can choose from various bed linen and pillows.

DE: Die Zimmer im Prague Mandarin Oriental bieten eine Fußbodenheizung und eine Auswahl an Bettwäsche und Kissen.

EN: The rooms at the Prague Mandarin Oriental feature underfloor heating , and guests can choose from various bed linen and pillows .

DE: Die Zimmer im Prague Mandarin Oriental bieten eine Fußbodenheizung und eine Auswahl an Bettwäsche und Kissen .

EN	DE
<blank> 1	<blank> 1
<unk> 2	<unk> 2
<s> 3	<s> 3
</s> 4	</s> 4
a 5	und 5
and 6	sie 6
the 7	mit 7
is 8	einen 8
with 9	der 9
in 10	ein 10

Tokenized text represented as bag of words vector based on vocabulary ids

2. Model Architecture: Approx. 220 Million Parameters

Model	
Model Type	seq2seq
Input Embedding Dimension	1,000
RNN Type	LSTM
# of hidden layers	4
Hidden Layer Dimension	1,000
Attention Mechanism	Global Attention

Umfasst   wifi   .

Includes   Wifi   .   Umfasst   wifi

Booking.com
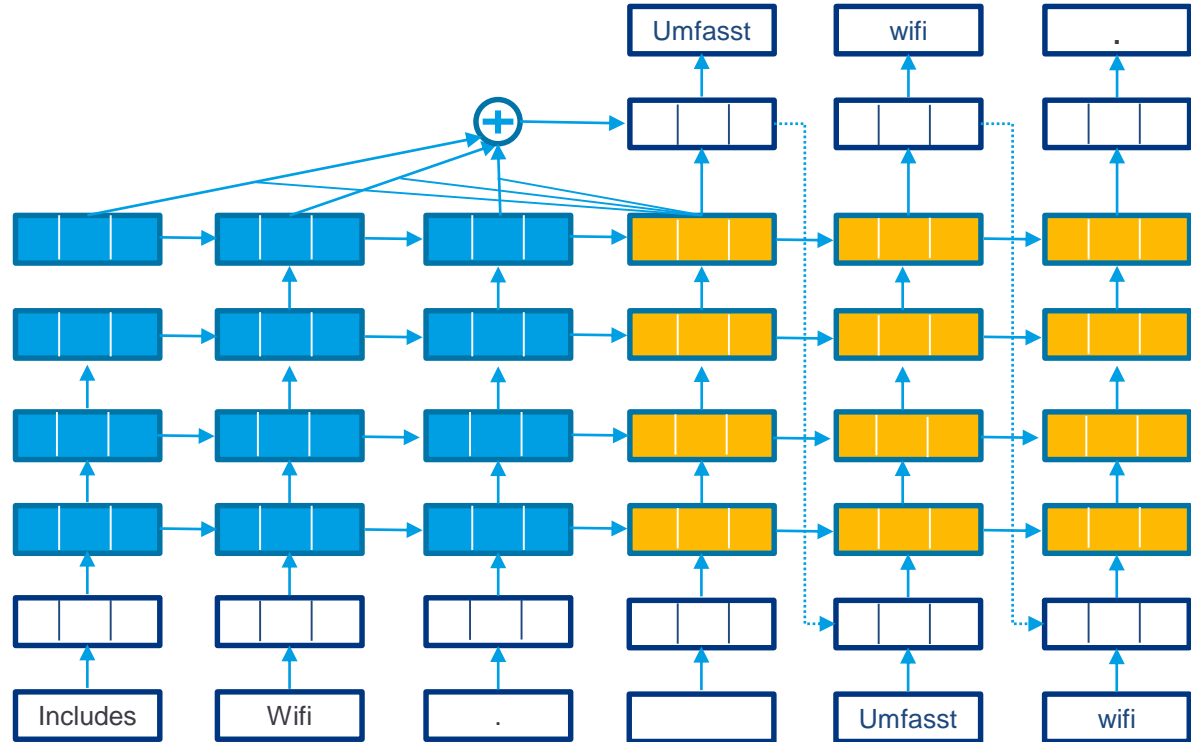
3. Training: 1 Epoch takes approx. 2 days on a single NVIDIA Tesla K80 GPU

Training	
Optimization Method	Stochastic Gradient Descent
Initial Learning Rate	1
Decay Rate	0.5
Decay Strategy	Decrease in Validation Perplexity <=0
Number of Epochs	5 - 13
Stopping Criteria	BLEU + sensitive sentences +constraints
Dropout	0.3
Batch Size	250

English -> German   English -> French

Perplexity Development

BLEU Score Development

Stopping Criteria: Sensitive Sentence Example

The neighborhood is very nice and safe
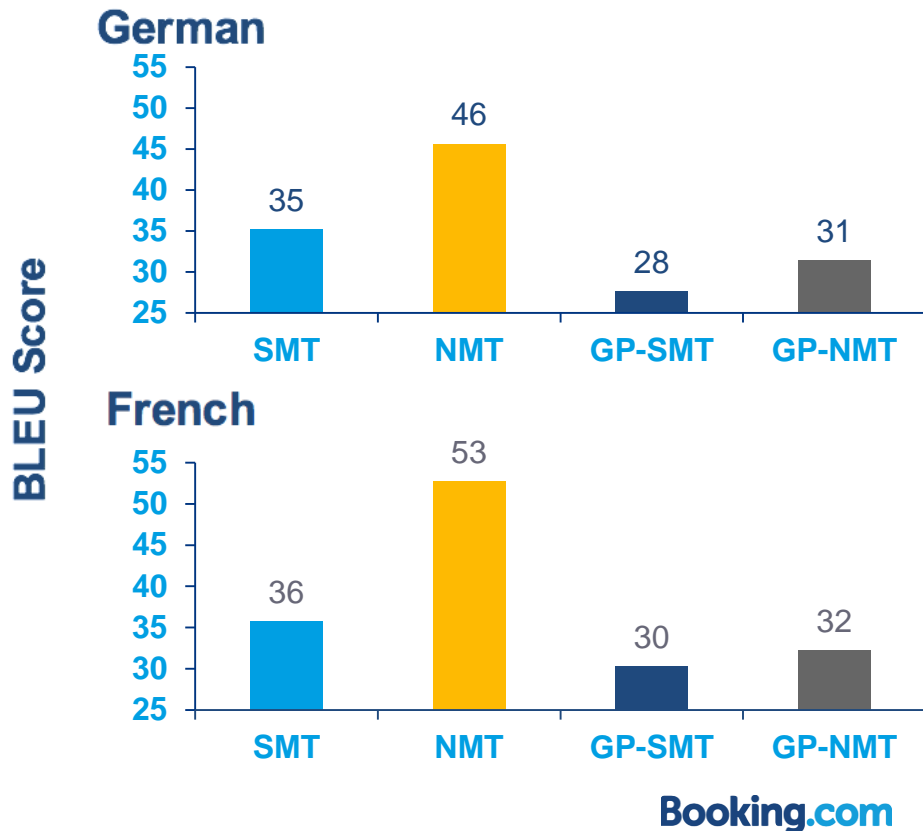
There is a safe installed in this very nice neighborhood

Booking.com

4. Translate: Unknown Word Handling

Translate	
Beam Size	30
Unknown Words Handling	Source with Highest Attention

	Good Example	**Bad Example**
Source	Offering a restaurant, Hodor Eco-lodge is located in Winterfell.	Free access to The Game entertainment Centre
Human Translation	Das Hodor Eco-Lodge begrüßt Sie in Winterfell mit einem Restaurant.	Kostenfreier Zugang zum Unterhaltungszentrum The Game
Raw Output	Das \<unk>\<unk> in \<unk> bietet ein Restaurant.	Kostenfreier Zugang zum \<unk>
Output with \<unk> replaced	Das Hodor Eco-lodge in Winterfell bietet ein Restaurant.	Kostenfreier Zugang zum Centre

5. Evaluate: Auto, Human, Length Analysis & A/B Tests

Evaluate	
Auto	BLEU WER
Human	A/F
Other	Length A/B Test

BLEU

# of words shared between MT output and human reference
- ➢ Benefits sequential words
- ➢ Penalizes short translations

WER

Variation of the word-level Levenshtein distance
- ➢ Measures the distance by counting insertions, deletions & substitutions

A/F Framework

- ➢ 3 evaluators per language
- ➢ Provided with original text and MT hypotheses, including human reference
- ➢ Not aware which system produced which hypothesis
- ➢ Asked to assess the quality of 150 random sentences from test corpus
- ➢ 4 level scale to both Adequacy & Fluency

Example:

Minor Mistake:
- EN: "there is a parking area available"
- DE: "es ist eine Garage verfügbar"

Major Mistake:
- EN: "there is a parking area available"
- DE: "es ist eine Aufbewahrungsstelle verfügbar"

Evaluation Results 1/5:
BLEU Score for German & French

German
BLEU Score

SMT 35 | NMT 46 | GP-SMT 28 | GP-NMT 31

French

SMT 36 | NMT 53 | GP-SMT 30 | GP-NMT 32

Booking.com

Our In-domain NMT system significantly outperforms all other MT engines

Both Neural systems consistently outperform their Statistical counter-parts

In-domain SMT beats General Purpose NMT

Compared to German, French improved much more from SMT to NMT

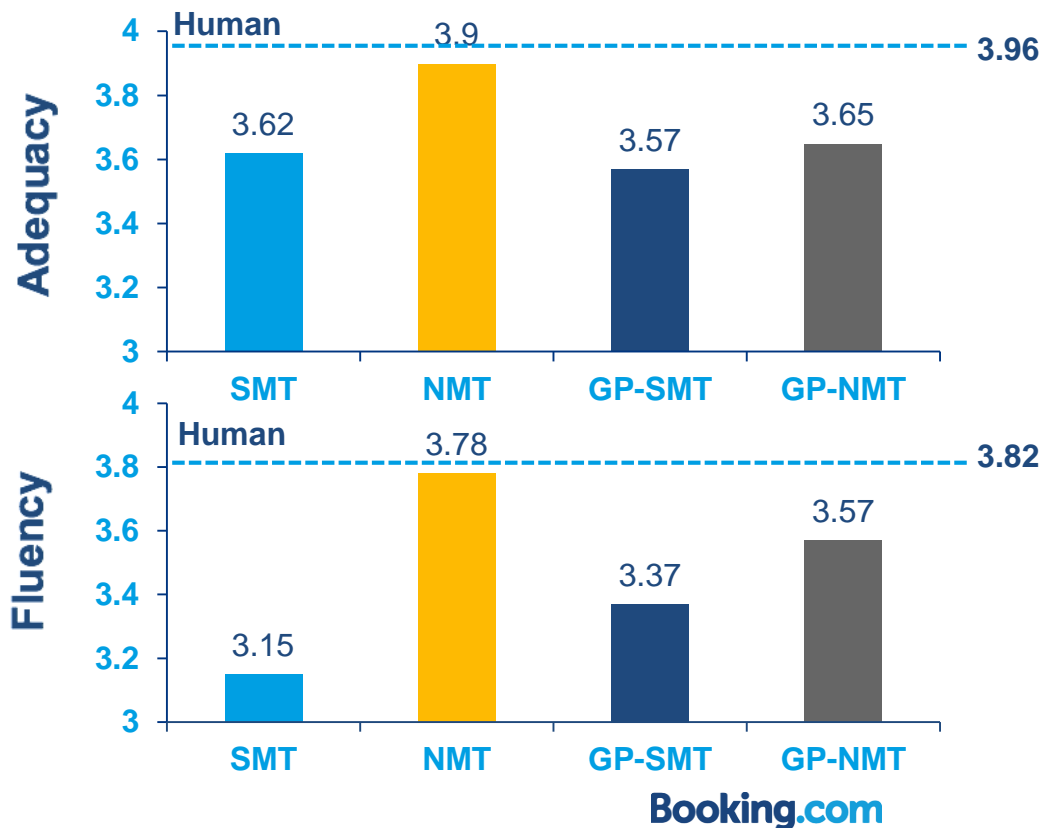Evaluation Results 2/5:
Adequacy/Fluency Scores for German

- Our In-domain NMT system still outperforms all other MT engines
- Both Neural systems still consistently outperform their statistical counter-parts
- However General Purpose NMT now beats In-domain SMT
- Particularly fluency score of our NMT engine is close to human level

Evaluation Results 3/5:
Adequacy/Fluency Scores for French
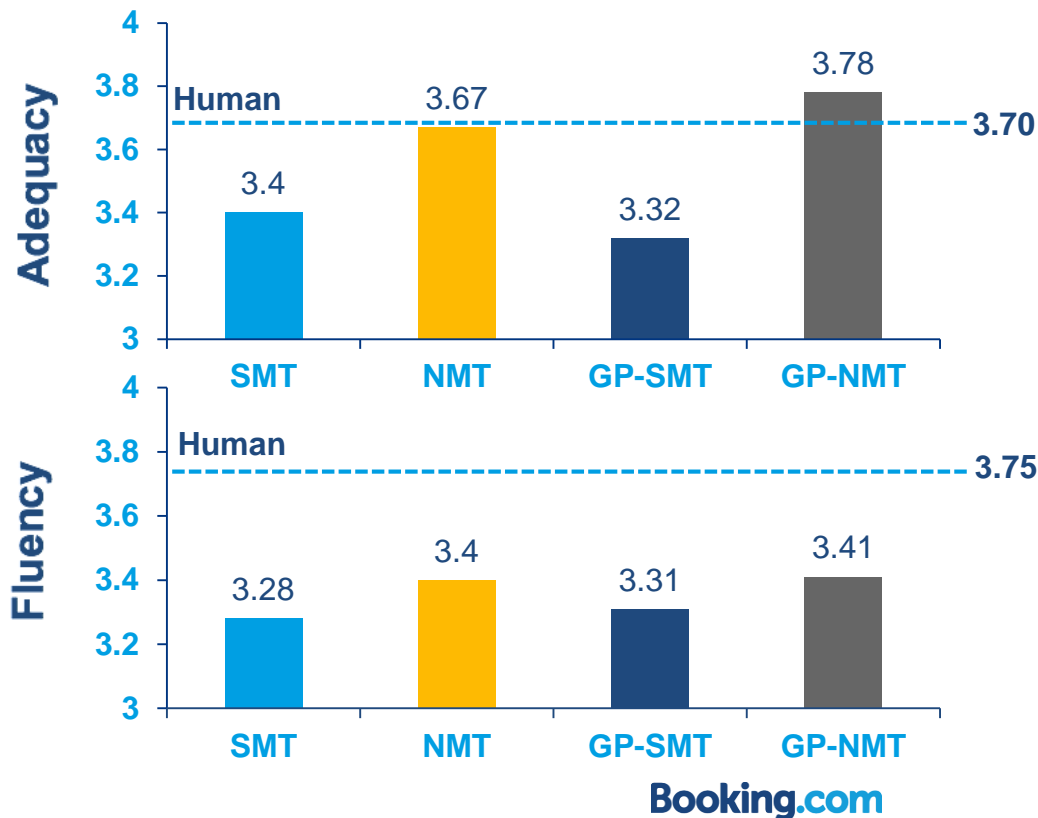
General Purpose NMT system outperforms others; conflicts with BLEU

Apparently General Purpose NMT even outperforms human level

Adequacy of both neural engines is almost at human level; fluency very far though

Compared to German, A/F scores are relatively less for French; conflicts with BLEU

Evaluation Results 4/5:
BLEU by Sentence Length for German and French
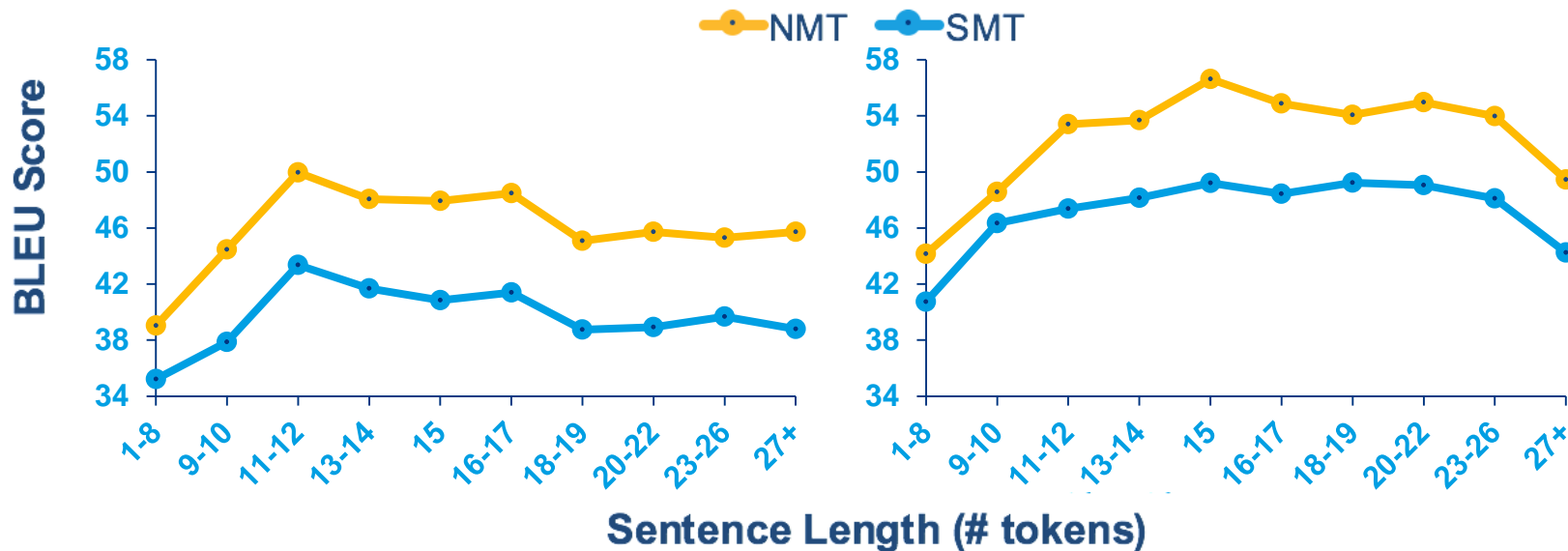
For longer sentences, though performance degraded, NMT still outperformed SMT

Initially performance increases with length, but reaches a peak soon & starts to decline then

For sentences longer than 27 tokens, NMT quality degrades faster than SMT

Evaluation Results 5/5:
Minus WER by Sentence Length for German and French

For longer sentences, though performance degraded, NMT still outperformed SMT

Initially performance increases with length, but reaches a peak soon & starts to decline then

For sentences longer than 27 tokens, NMT quality degrades faster than SMT

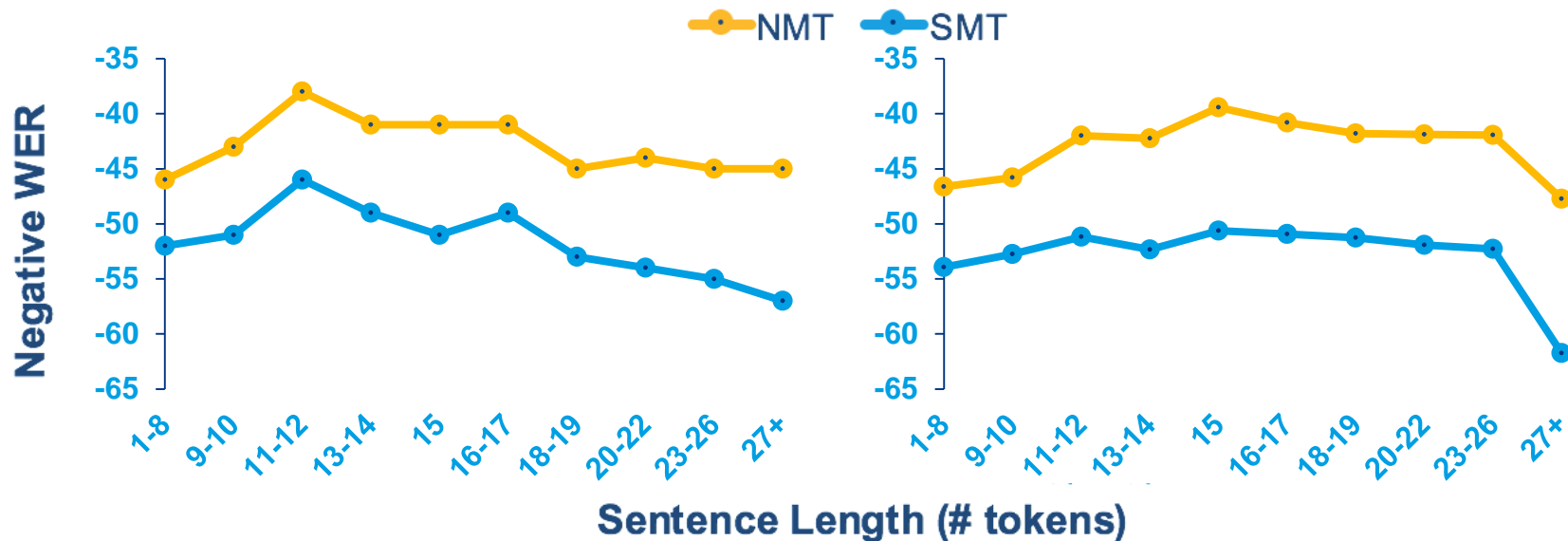# A/B Tests to validate the hypothesis that MT'ed hotel descriptions have higher conversion than no translation

German visitor

50% see base with no translation

50% see variant with machine translation

Offering free WiFi and a garden, VSG Apartment Petrska is situated in Prague, 900 metres from Old Town Square. Prague Astronomical Clock is 1 km away.

The accommodation comes with a seating and dining area. All units feature a kitchen equipped with a dishwasher and microwave. A fridge and coffee machine are also provided. Towels and bed linen are offered.

Wenceslas Square is 1.2 km from VSG Apartment Petrska. The nearest airport is Vaclav Havel Prague Airport, 12 km from the property.

Mit kostenfreiem WLAN und einem Garten erwartet Sie das VSG Apartment Petrska in Prag, 900 m vom Altstädter Ring entfernt. Die Astronomische Uhr von Prag erreichen Sie nach 1 km.

Die Unterkunft verfügt über einen Sitz- und Essbereich. Alle Unterkünfte verfügen über eine Küche mit einem Geschirrspüler und einer Mikrowelle. Ein Kühlschrank und eine Kaffeemaschine sind ebenfalls vorhanden. Handtücher und Bettwäsche werden gestellt.

Der Wenzelsplatz liegt 1.2 km vom VSG Apartment Petrska entfernt. Der nächste Flughafen ist der 12 km von der Unterkunft entfernte Flughafen Prag.

Booking.com

Few Interesting Examples from French translations

	Source	Translation
Good Word Sense Disambiguation	The neighbourhood is very safe. There is a safe installed in the room.	Le quartier est très sûr. Vous trouverez un coffre-fort dans la chambre.
Bad Out of domain sentence	The owners are super right wing.	Les propriétaires se trouvent dans une aile droite.
Ugly OOV words	Sdfdlsfsldk offers free breakfast	Le offers sert un petit-déjeuner gratuit.

Booking.com

Conclusion & Future Work

Conclusion.

- NMT consistently outperforms SMT

- In case of German, in-house NMT is also better than online general purpose engines in our application

- Fluency of NMT is close to human translation level

- In our application the relative performance of NMT against SMT does not degrade with increased sentence length

Future Work.

- **Improve Unknown Word Handling**
  - Explore open vocabulary techniques for UNK handling; sub-word tokenization using byte pair encoding

- **Identify Business Sensitive Translation Errors**
  - 'free' being translated to 'available'

- **Expand to other languages for hotel descriptions**
  - Particularly Asian languages like Chinese, Japanese etc.

- **Expand to other use cases for MT at Booking.com**
  - User generated content like customer reviews, messages etc.

Booking.com

Thank You
Questions?

We're hiring!
workingatbooking.com

Nishikant Dhanuka from Booking.com
linkedin.com/in/nishikantdhanuka/