The STARway to MT heaven
Uniting human & artificial intelligence

Judith Klein
STAR Language Technology & Solutions

TC40 – ASLING 2018
MT heaven?

Google's AI translation system is approaching human-level accuracy, but there's still significant work to be done.

"Nearly Indistinguishable From Human Translation"—Google Claims Breakthrough

By Florian Faes on September 27, 2016

DeepL schools other online translators with clever machine learning

A major breakthrough in the field of Google translation integration: translation quality is close to that of humans.

Google's New Service Translates Languages Almost as Well as Humans Can

A jump in the fluency of Google's language software will help efforts to make chatbots less lame.

by Tom Simonite | September 27, 2016

DeepL: A Modern-day David and Goliath Story

Posted on 7 May 2018 by Margot BarthelEMY | Author: Margot BarthelEMY | Machine translation
STAR Group (since 1984)

Translation Services & Translation Technology
▲ Transit 2.7 & Logos

late 1990s......
STAR’s Translation Strategy

▲ Respond to customer requests

▲ Develop translation technology for translators

... for successful translation business
▲ …Transit NXT & many MT systems
▲ Transit NXT & many MT systems

▲ STAR MT
  • SMT 2015
  • NMT 2018
MT integration in Transit NXT

Pretranslation

STAR Transit

TM

TM

Fuzzy matches

Project

Translation

Dictionaries

Source text

Translated text
MT integration in Transit NXT

- Pretranslation
  - STAR Transit
  - Project + MT

- Translation
  - Fuzzy matches
  - MT suggestions

- STAR MT
- SMT/NMT

- Not pretranslated

- Source text
- Dictionaries
- Translated text
MT integration in Transit NXT

Not pretranslated

Pretranslation

STAR MT

SMT/NMT

Online MT

Fuzzy matches
MT suggestions

Source text

Dictionaries

Translated text
MT for translators

▲ MT not inserted automatically
MT for translators

▲ MT not inserted automatically

▲ Additional source
  • instead of no match
MT for translators

▲ MT not inserted automatically

▲ Additional source
  • instead of no match
  • in addition to fuzzy match
MT for translators

▲ MT not inserted automatically

▲ Additional source
• instead of no match
• in addition to fuzzy match

▲ Clearly labelled

<table>
<thead>
<tr>
<th>Source</th>
<th>Fuzzy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ref</td>
<td>they know there is a better way to communicate.</td>
</tr>
<tr>
<td>New</td>
<td>they know there's a more efficient way to communicate.</td>
</tr>
<tr>
<td>Ref</td>
<td>sie kennen einen besseren Weg der Kommunikation.</td>
</tr>
<tr>
<td>New</td>
<td>Sie wissen, dass es eine höhere Methode ist, um miteinander zu kommunizieren.</td>
</tr>
</tbody>
</table>

Machine Translation (Import MT)
MT for translators

▲ MT not inserted automatically

▲ Additional source
  • instead of no match
  • in addition to fuzzy match

▲ Clearly labelled

▲ Unknowns indicated
MT for translators

▲ MT not inserted automatically

▲ Additional source
  • instead of no match
  • in addition to fuzzy match

▲ Clearly labelled

▲ Unknowns indicated

▲ TM-validated MT
MT for translators – MT position

List MT suggestions in the fuzzy window

- Above the best fuzzy match
- In the same way as a fuzzy match at (%): 75%

<table>
<thead>
<tr>
<th>Machine Translation (Import MT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implausible data were received from control unit “Traction system”</td>
</tr>
<tr>
<td>Wurden nicht plausible Daten empfangen von Steuergerät Traktionssystem</td>
</tr>
</tbody>
</table>

- 79% D:\Transit NXT\projects\MT_Example\AEXTR_1
- Implausible data were received from control unit “Battery”.
- Implausible data were received from control unit “Traction system”.
- Vom Steuergerät “Batterie” wurden nicht plausible Daten empfangen.

- 74% D:\Transit NXT\projects\MT_Example\AEXTR_1
- Implausible data were received from control unit “Radar sensorics”.
- Implausible data were received from control unit “Traction system”.
- Vom Steuergerät “Radarsensoren” wurden nicht plausible Daten empfangen.
MT for translators – MT position

<table>
<thead>
<tr>
<th>List MT suggestions in the fuzzy window</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Above the best fuzzy match</td>
</tr>
<tr>
<td>☐ In the same way as a fuzzy match at (%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>List MT suggestions in the fuzzy window</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Above the best fuzzy match</td>
</tr>
<tr>
<td>☐ In the same way as a fuzzy match at (%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MT suggestions in the fuzzy window</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine Translation (Import MT)</td>
</tr>
<tr>
<td>Implausible data were received from control unit “Traction system”</td>
</tr>
<tr>
<td>Implausible data were received from control unit “Battery”</td>
</tr>
<tr>
<td>Implausible data were received from control unit “Radar sensorics”</td>
</tr>
<tr>
<td>Implausible data were received from control unit “Traction system”</td>
</tr>
<tr>
<td>Implausible data were received from control unit “Battery”</td>
</tr>
<tr>
<td>Implausible data were received from control unit “Radar sensorics”</td>
</tr>
<tr>
<td>Implausible data were received from control unit “Traction system”</td>
</tr>
<tr>
<td>“Implausible data were received from control unit “Battery”.”</td>
</tr>
<tr>
<td>“Implausible data were received from control unit “Traction system”.”</td>
</tr>
<tr>
<td>“Implausible data were received from control unit “Radar sensorics”.”</td>
</tr>
<tr>
<td>“Implausible data were received from control unit “Traction system”.”</td>
</tr>
<tr>
<td>“Implausible data were received from control unit “Battery”.”</td>
</tr>
<tr>
<td>“Implausible data were received from control unit “Radar sensorics”.”</td>
</tr>
<tr>
<td>“Implausible data were received from control unit “Traction system”.”</td>
</tr>
<tr>
<td>“Vom Steuergerät “Batterie” wurden nicht plausible Daten empfangen.”</td>
</tr>
<tr>
<td>“Vom Steuergerät “Radarsensoren” wurden nicht plausible Daten empfangen.”</td>
</tr>
</tbody>
</table>

© STAR Group
Uniting human and artificial intelligence

Transit project without MT

- No match: 28%
- Low Fuzzies: 10%
- Medium Fuzzies: 15%
- High Fuzzies: 9%
- Perfect Matches: 38%

Transit projects with STAR MT

- No match: 22%
- Low Fuzzies: 8%
- Medium Fuzzies: 13%
- High Fuzzies: 14%
- Perfect Matches: 43%

Average 20% more for no match segments
Thank you

Judith Klein
STAR Language Technology & Solutions

TC40 – ASLING 2018