What? Why? How?
Factors that impact the success of commercial MT projects

John Tinsley
Iconic Translation Machines
“Why would I need MT?”

What’s the MT value proposition?

**Why MT?**

- Speed
- Cost savings
- Time to market
- Your competitors are doing it!

**Why Now?**

- Volume of content is growing
- Demand, more words less time
- Growth facilitator
- #FOMO – you’re missing out on business
“What type of MT is it?”

- statistical
- hybrid
- rule based
- syntactic
- custom
- intelligent
- hierarchical
- composite
- cloud
The goal of the MT here is to be good enough so that - on the whole – with TMs, translators are faster post-editing some segments

Challenges

- development has to focus on reducing needs for edits, not necessarily anything else
- translator acceptance always a big barrier
- evaluation can take time and has many factors
- pricing models
The goal is to produce MT that’s fit for a particular purpose as is.

Arguably easier from an MT development perspective.

Often high-volumes = more achievable.
The 8 Factors influencing MT suitability

- **TM Leverage**
  - High TM Leverage
  - Low MT Effectiveness

- **Quality Required**

- **Buyer Maturity**

- **Integration Needs**

- **Training Data**

- **Language**
  - French
  - Spanish
  - Hungarian
  - Portuguese
  - Basque

- **Volume**

- **Content**

- **Reality-Check**

---


Austin, Oct 28 - Nov 1, 2016 | p. 292
Not all languages are created equal
The more words...the better...the worse?
The bed was two twin beds put together and me and my girlfriend kept falling in the middle (since we like to cuddle) and that was irritating. Late night room service was awesome.
The more experience the LSP has with onboarding/training vendors, and the more experience the vendor has with MT, the more feasible the adoption of MT will be.
Standard vs Custom Integration

“instant” solution costs rise proportionality with the number of languages and the throughput needs
High TM Leverage

<table>
<thead>
<tr>
<th>Matches</th>
<th># words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>403,803</td>
</tr>
<tr>
<td>100%</td>
<td>585,459</td>
</tr>
<tr>
<td>95-99%</td>
<td>50,366</td>
</tr>
<tr>
<td>85-94%</td>
<td>41,604</td>
</tr>
<tr>
<td>75-84%</td>
<td>32,319</td>
</tr>
<tr>
<td>50-74%</td>
<td>18,972</td>
</tr>
<tr>
<td>No Match</td>
<td>81,119</td>
</tr>
<tr>
<td>Total</td>
<td>1,213,643</td>
</tr>
</tbody>
</table>

Only 8% of all words go to MT

Low MT Effectiveness
Quality requirements

• Fully automatic human quality

• 300% post-editing productivity

• French to Spanish == English to Korean

• Best performance out of the box
The 8 Factors influencing MT suitability

- TM Leverage
- Quality Required
- Integration Needs
- Buyer Maturity
- Training Data
- Language
- Content
- Volume
What questions should YOU be asking?

- “What volume of words do you estimate for the project?”
- “Do we have translation memories, glossaries that are relevant? Can we create them?”
- “If so, what leverage are we getting?”
- “To we have post-editors? Access to a supply chain?”
  - “What experience do they have?”
- “Where will MT fit in the workflow (depending on the use case)?”
- “What variety is there in the content that the MT will be processing?”
- “Why aren’t you using Google Translate?”
- “Is there sufficient budget for this project?”
How much training data do I need?

How frequently can I retrain the engine?

What happens to my data?

Do you do language X?

How good is the quality?

How do you measure performance over time?

What happens to my data?