Understanding Challenges to Enterprise MT Adoption

Bart Mączyński, VP Machine Learning, SDL
Agenda

- Introduction
- New Buyers, New Misconceptions
- The Challenges
  - The Use Case Challenge
  - The Technical Challenge
  - The Linguistic Challenge
- What’s Next
Introduction

Bart Mączyński

- VP Machine Learning, Solutions Consulting
- Expertise in translation management, TM, MT, terminology systems
- Over 20 years of experience in the field
- Focus on commercial applications of Linguistic AI
Background

**Rise of the Machines**

- MT is now a viable solution for the enterprise
- Recent advancement opened up new MT use cases
- MT is now directly exposed to new buyer communities
- These new buyers may not have much experience in translation management
New Buyers, New Misconceptions

AI, ML, MT Hype

- “MT replaces human translators”
- “MT can learn from what it translates”
- “MT can handle all content”
- “MT quality is amazing across the board”
- “MT is cheap”
- “Anyone can build an MT system”
- “I’ve read about GPT-3, all my content issues will be solved soon”
The Three Challenges
The Use Case Challenge
The Use Case Challenge

• Understanding if, and how, MT can be utilized
• Providing a pathway to the most optimal translation option

Things to Consider

• Language pair coverage
• Quality
• Volume
• Speed
• Security
On-premise vs. Cloud
The Technical Challenge
The Technical Challenge

• Understanding how the MT solution needs to be deployed
• Providing a pathway to acceptable TCO

Things to Consider

• Number of LPs and individual LP scalability
• Translation speed and latency
• Cost of achieving scale (hardware, hosting)
• Flexible licensing model (e.g. for seasonal peaks)
• Integrations and burden of maintenance
• Security and business continuity
The Linguistic Challenge

• Understanding the current limits of MT
• Providing a pathway to sufficient quality

Things to Consider

• MT still comes out as too direct and non-idiomatic
• High level of fluency can mask other issues
• Humans excel at detecting extra-textual context
• Even commercial-grade MT systems struggle with messy inputs
• Bias in training data is reflected in MT output
Example 1 – Fluent Nonsense

Phototherapy exposure in this range is used in the treatment of hyperbilirubinemia in newborn children.

Fototerapia ekspozycji w tym zakresie jest stosowana w leczeniu hiperbilirubinemii u noworodków.

fototerapia ekspozycji = phototherapy of exposure
Example 2 – Cross Domain Issues

System 1

Administration should be performed by an individual who has been adequately trained in injection techniques.

Podawanie powinno być wykonywane przez osobę, która została odpowiednio przeszkolona w zakresie technik iniekcji.
Example 2 – Cross Domain Issues

System 2

Administration should be performed by an individual who has been adequately trained in injection techniques.

Podawanie leku powinno być wykonywane przez osobę, która została odpowiednio przeszkolona w zakresie technik wstrzykiwania.
Example 2 – Cross Domain Issues

System 1

System administration should be performed by an individual who has been adequately trained.

Podawanie systemu powinno być wykonywane przez osobę odpowiednio przeszkoloną.
Example 2 – Cross Domain Issues

System 2

System administration should be performed by an individual who has been adequately trained.

Administrowanie systemem powinno być wykonywane przez osobę, która została odpowiednio przeszkolona.
Example 3 – Gender Bias

The doctor went home.
The nurse went home.
The professor went home.
The cleaner went home.

[m] Der Arzt ging nach Hause. Доктор пошел домой.
[m] Der Professor ging nach Hause. Профессор пошел домой.
[f] Die Putzfrau ging nach Hause. Уборщица пошла домой.
Example 4 – Imperfect Inputs

This is a test of the emergency alert system.

To jest test systemu alarmowegh.

This is a test of the emergency alert system

To jest test systemu alarmowegh w sytuacjach awaryjnych
Example 5 – Missing Context

Wymiana puszek i instalacji w starym domu.
Replacement of cans and installations in an old house.

Replacement of junction boxes and wiring in an old house.