Languages and knowledge: AI and machine translation in scientific communication

Maria Teresa Zanola
CEL/ELC - REALITER

AI4EI Conference
Linguistic Rights and Language Varieties in Europe in the Age of AI
Turin, 23/04/2021
Introduction

- Terminology contributes to the dissemination of knowledge, developing profound relationships with the socio-economic, technical and scientific evolution of the relative referents.
- It promotes the dissemination of innovation, grafting it onto established linguistic heritage and inserting it within a precise conceptual network.
Terminology...

is everyone’s heritage and expressive wealth

offers its tools to specialists, translators, scientific journalists, technical editors as well as all professional categories and ordinary citizens

is the vehicle, in every trade and profession, be it scientific, humanist, institutional, legal, economic, financial, technical or operational, for transmitting and expressing all innovations within our conceptual heritage

ai4ei Conference - Turin, 23/04/2021
The more conscious terminology is in its linguistic, conceptual and cultural components, the greater its degree of communicative effectiveness.
Terminology makes it possible to reconstruct the life of knowledge of all kinds.

The importance of terminology in communicating science and technology, as well as in transmitting knowledge in the arts and crafts.
How can we preserve this terminological heritage - based on languages and knowledge - in scientific communication using machine translation?
Are there any disadvantages?
Are there any advantages?

We will be offering various reflections on the subject, from a multilingual perspective.
Terminology and automatic translation: disadvantages and advantages

Studies in the field

TM and Multilingualism
The disadvantages

• The lack of terminological accuracy

• Linguistic errors:
  • aspect-related nuance
  • changes in word order
  • structural changes; inversion of dependency relations

• The undermining of the capacity for autonomous processing of translations
The advantages

• The possibility of accessing scientific information in several languages
• Enhancing the value of all languages
• Achievements in technical communication are now happening in scientific communication
A survey of UCSC students: their perception about TM

The analysis of a target corpus of TM technical and scientific texts (FR/EN/DE)

ai4ei Conference - Turin, 23/04/2021
Machine translation (MT) technologies have improved significantly over the last two decades, with developments in phrase-based statistical MT (SMT) and, more recently, neural MT (NMT).

However, most of these methods rely on the availability of ample parallel data for training the MT systems. As these resources are not available for the majority of language pairs, current technologies often fall short in their applicability to low-resource languages.

This technology is now used to overcome language barriers in high-risk settings such as hospitals and courts.

In its current state, MT technology can exacerbate social inequalities and put certain user communities at greater risk.

This issue merits further attention from researchers and policymakers.

“A deep-learning system that significantly out-performed professional-agency English-to-Czech news translation in preserving text meaning (translation adequacy).”

This work approaches the quality of human translation, even surpassing it in terms of adequacy in certain circumstances. This suggests that deep learning may have the potential to replace humans in applications where conservation of meaning is the primary aim.
Training translators in...
- machine translation: statistical, rule-based, example-based and hybrid
- machine-aided translation and interpreting
- the annotation, alignment and searchability of translation data
- multi-modal and audiovisual translation
• TM can enhance translation operability, quality and speed and meet the demands of translation practice, accelerating communication between different countries and different languages…
AI and automatic translation

Human translation
Safeguarding knowledge

Protecting the intellectual heritage represented by terminology in source languages
• “we took the trouble of going to their workshops to interview them, transcribing what they said, developing their thoughts, extracting the terms specific to their professions, organising them into tables and defining them, conversing with those who had supplied us with their memories….”

“Discours préliminaire”, *Encyclopédie, 1751*

To conclude…
Thank you