universität wien

Impact of Domain-Adapted Multilingual Neural Machine Translation in the Medical Domain

Miguel Rios, Raluca-Maria Chereji, Alina Secară, Dragoș Ciobanu

HAITrans (Human and Artificial Intelligence in Translation) research group

Centre for Translation Studies

University of Vienna

https://haitrans.univie.ac.at/

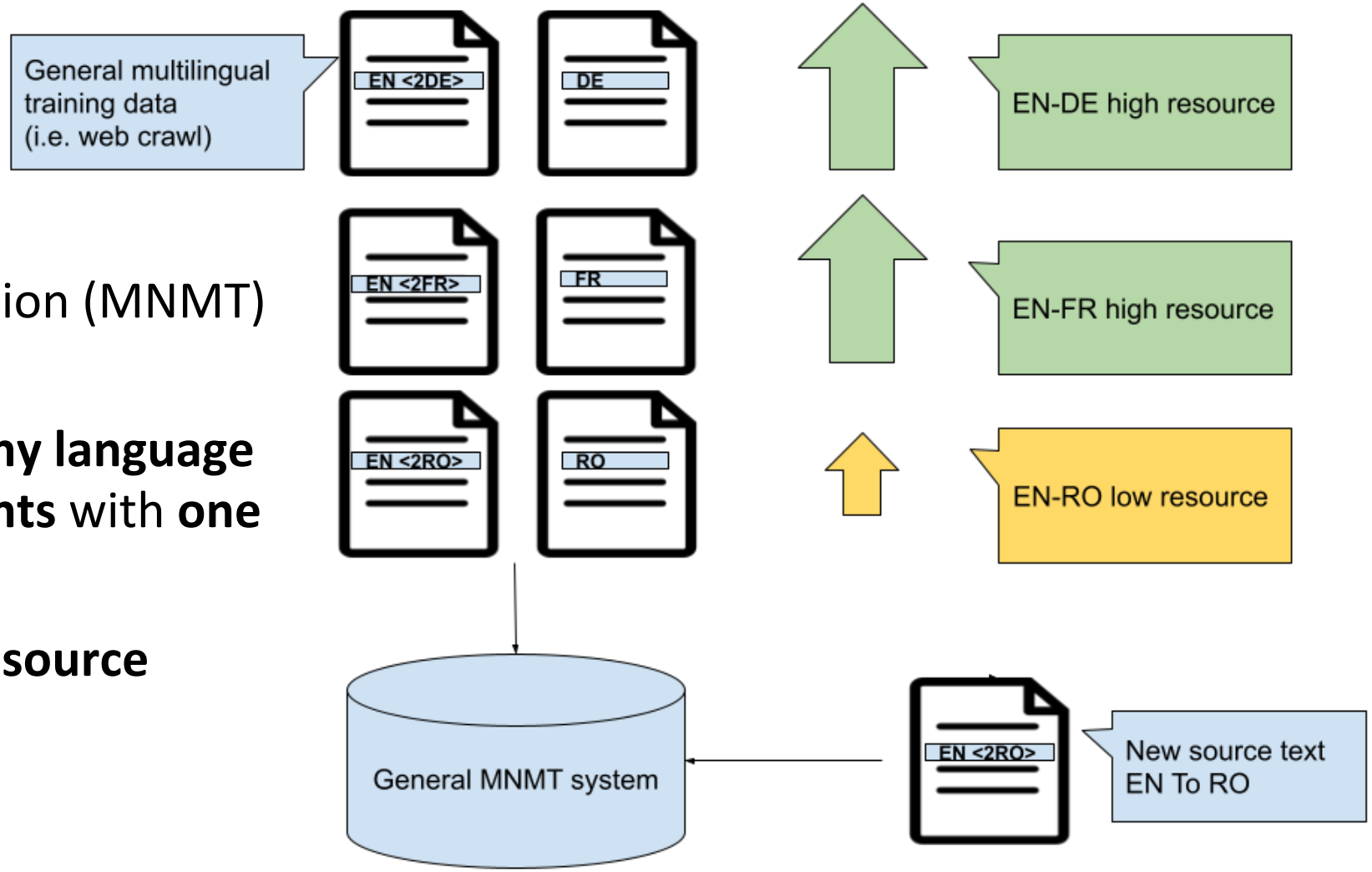
universität wien

Content

- Introduction
- Motivation
- Experiments
- Results
- Error Analysis
- Conclusions
- Future Work

Introduction

- Multilingual Machine Translation (MNMT) (Johnson *et al*., 2017)
  - MNMT models leverage **many language pairs** and **millions of segments** with **one translation system**.
  - Translation quality of **low-resource** languages **benefits** from the **high-resource** languages.



Introduction

ENGLISH

The stratosphere extends from about 10km to about 50km in altitude.

KOREAN

성층권은 고도 약 10km부터 약 50km까지 확장됩니다.

JAPANESE

成層圏は、高度 10km から 50km の範囲にあります.

EN-JA, EN-KO

From: (Johnson *et al.*, 2017)

universität wien

Motivation

- Medical MT is a **high-risk** and **low-resource** specialised domain.
- Challenges in Medical MT:
  - Accurate translation of **terminology**.
  - Typos, writing style, and code switching.
- Why **domain adaptation**?
  - Leveraging large open-source MNMT models instead of training from scratch.

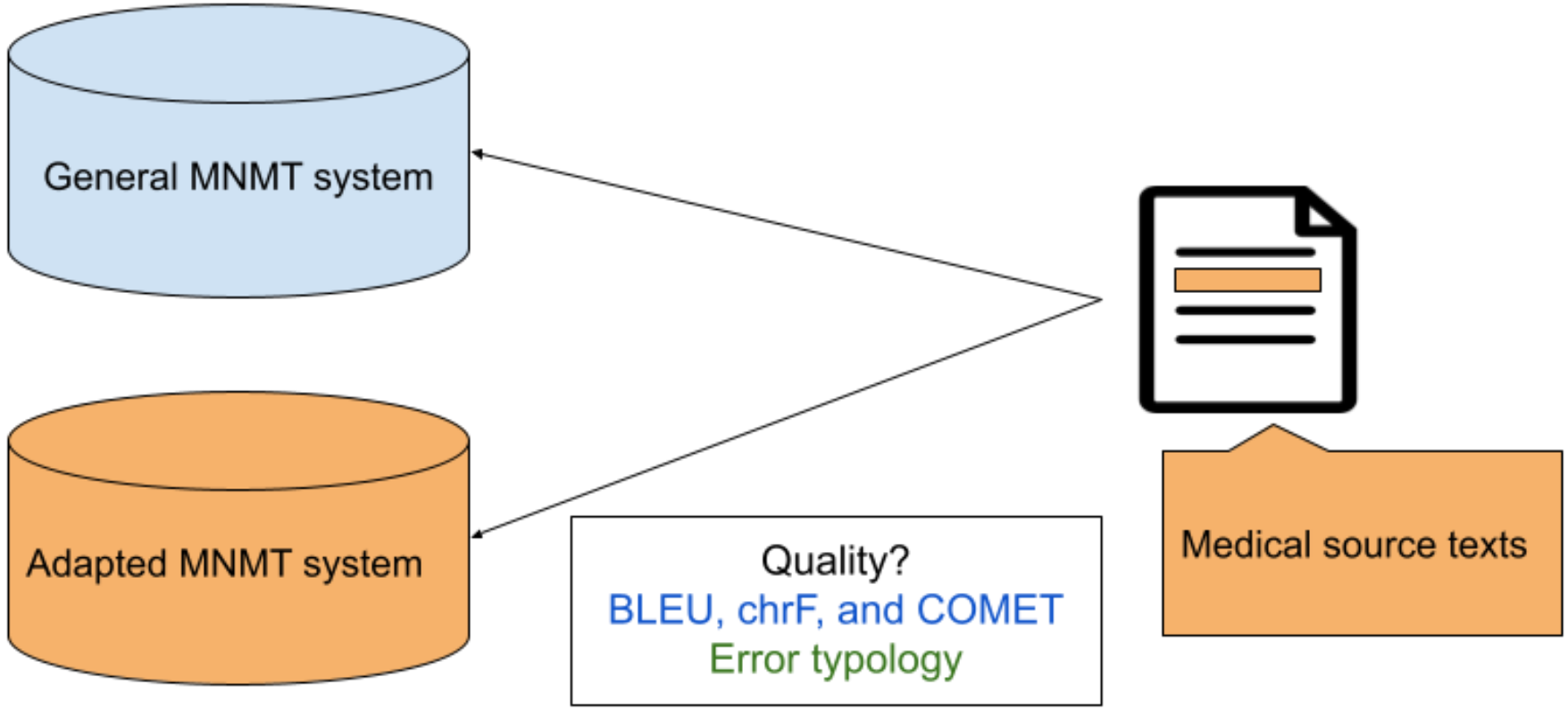
universität wien

Motivation (cont.)

- We study the quality of a **domain-adapted** MNMT in the **medical domain** for a low-resource language pair in terms of **automatic metrics** and an **error typology** for **terminology**.
  - **Automatic metrics** do not highlight errors in MT.
  - Expert annotation of **terminology** translation errors with an **error typology**.

universität wien

Motivation (cont.)



universität wien

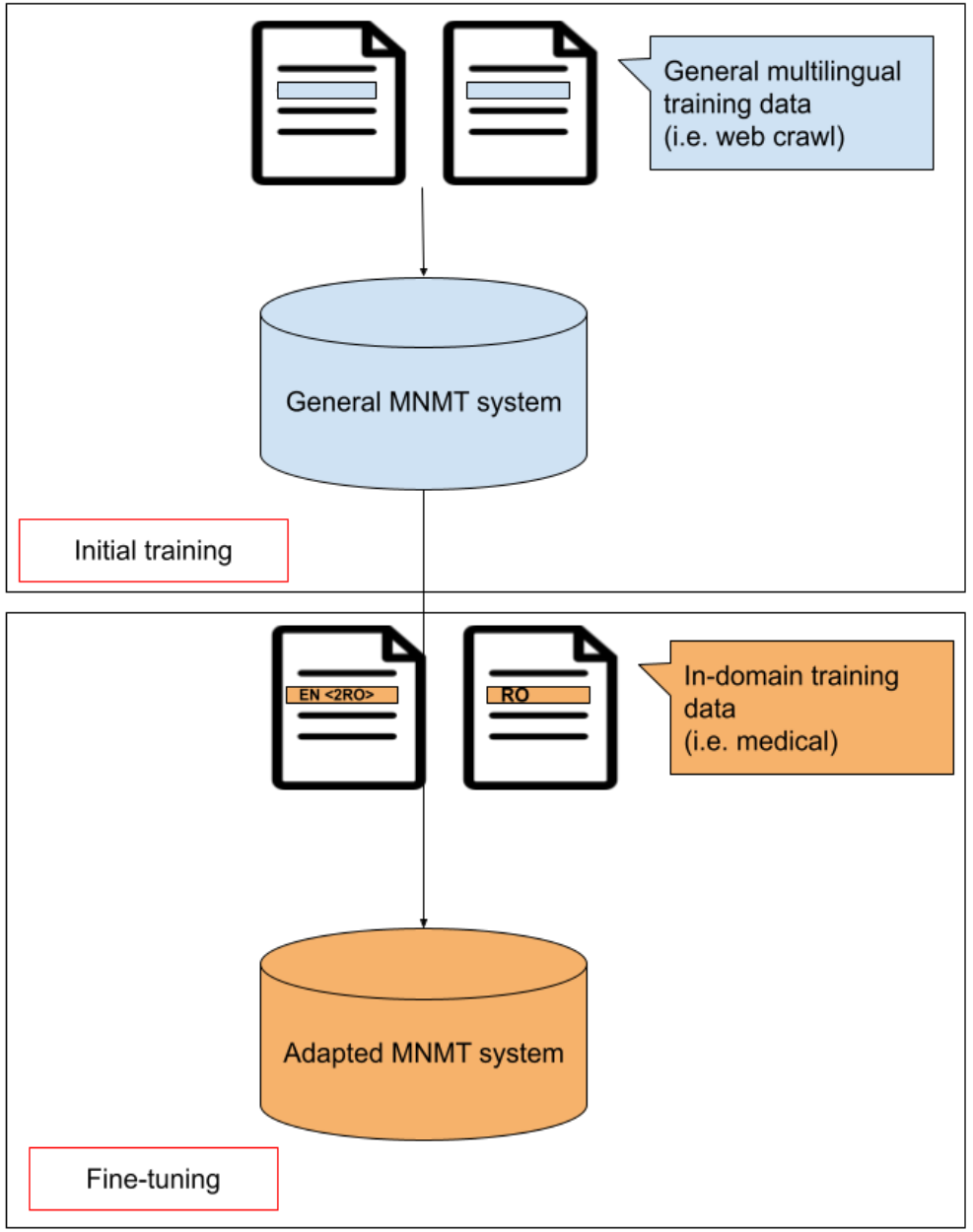
Experiments

- Data:
  - English-Romanian (EN-RO)
  - Training dataset: **EMEA** (European Medicines Agency) corpus with **775,904 training**, and **7,837 validation segments** (CLARIN:EL, 2015).
  - Test dataset: **Medline abstracts** from medical scientific publications with **291 segments**.

universität wien

Experiments (cont.)

- Models:
  - General MNMT: **MBart** (Liu *et al.*, 2020)
  - **Fine-tuned** MBart with EN-RO EMEA.



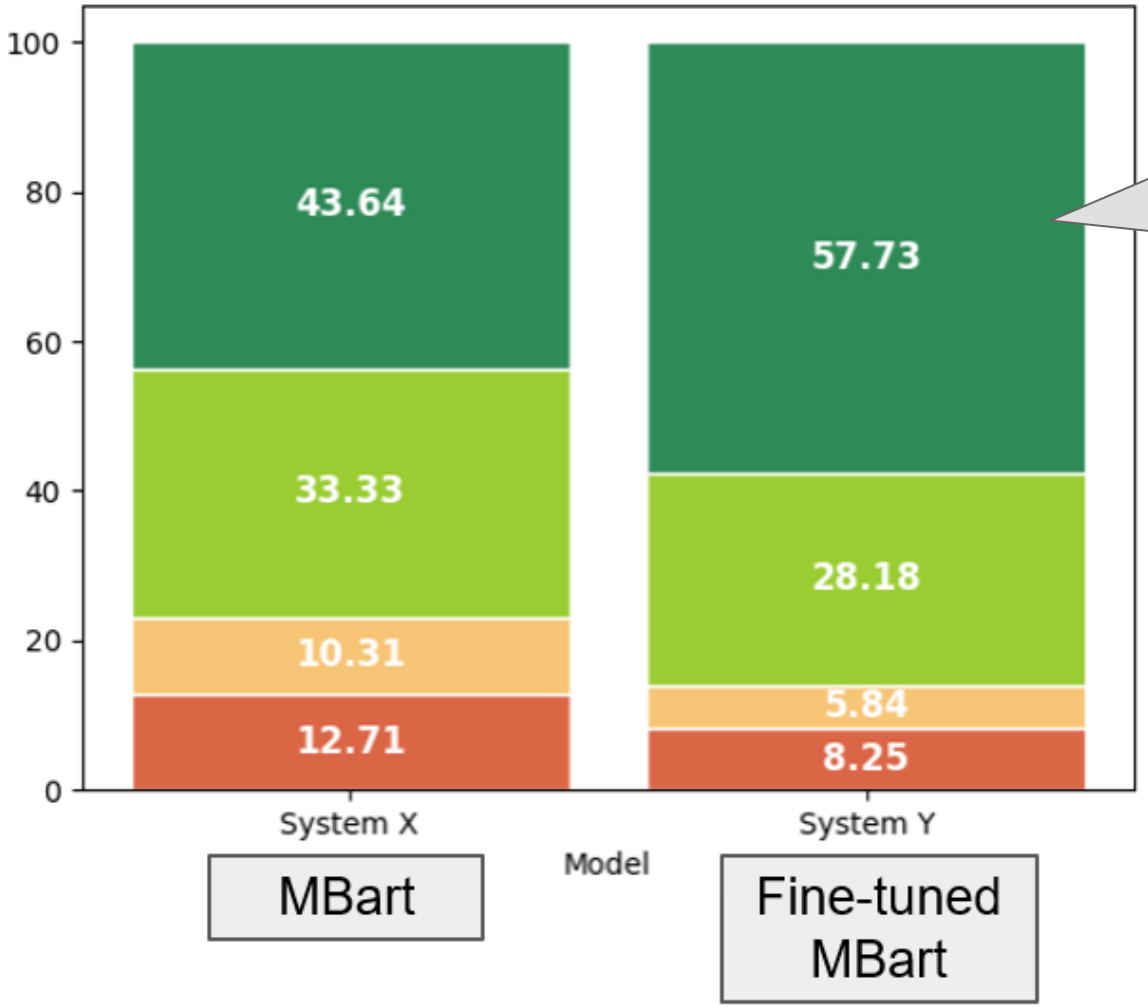
universität wien

Results with Automatic Metrics

	BLEU ↑	chrF ↑	COMET ↑
MBart	21.9	51.5	0.556
Fine-tuned MBart	**25.8****	**54.9**	**0.663**

BLEU is statistically significant p=0.001

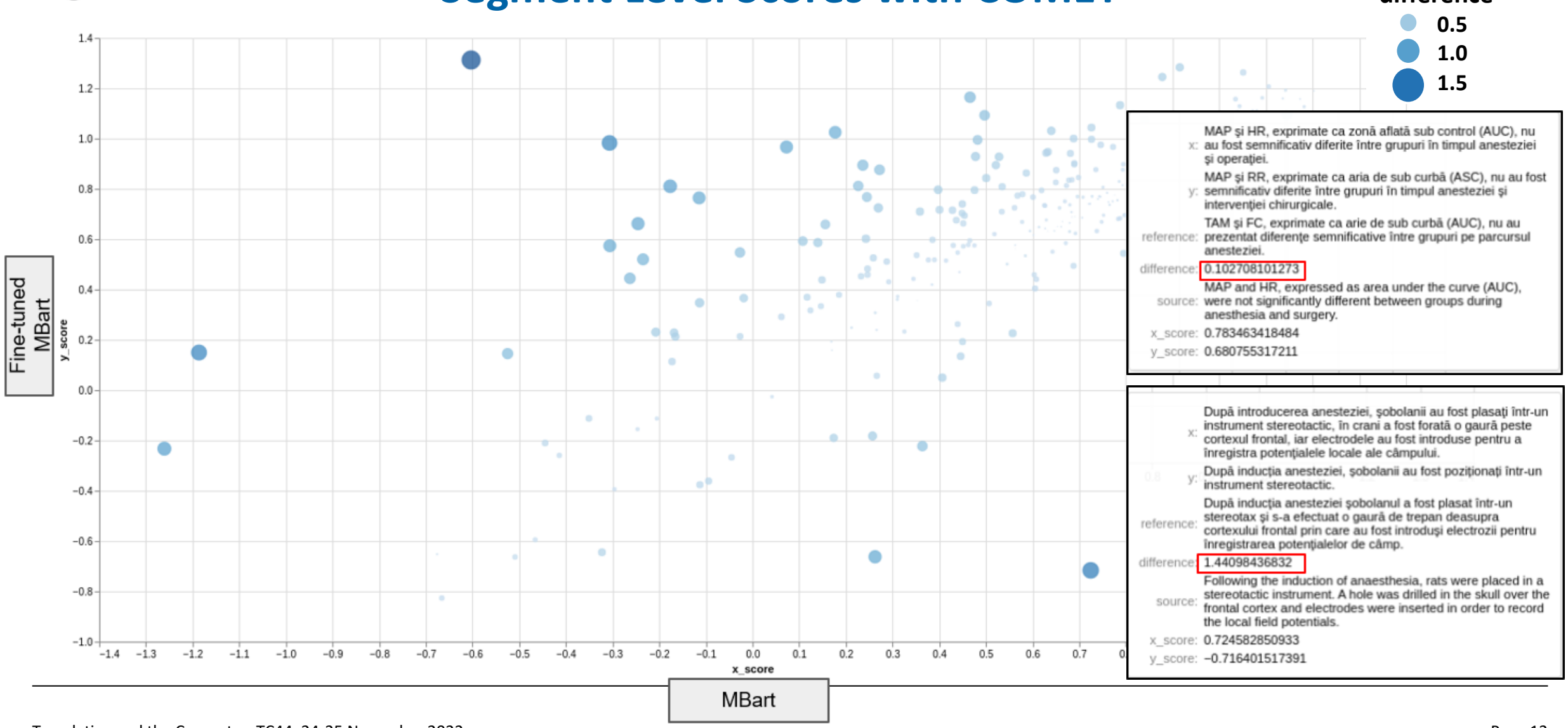
Segment Level Scores with COMET



Fine-tuned MBart has a bigger percentage of segments with a **high score** (meaning that they are close to the reference translation).

MT-Telescope:
https://github.com/unbabel/mt-telescope

Segment Level Scores with COMET



universität wien

Error Analysis

- **Error typology** with a focus on **terminology** for human evaluation (Haque *et al*., 2019)
- Eight error categories:
  - **(1)** Partial error, **(2)** Source term copied, **(3)** Inflectional error, **(4)** Reorder error, **(5)** Disambiguation issue in target, **(6)** Incorrect lexical selection, **(7)** Term drop, **(8)** Other error
- Sample of **12 abstracts** with a total of **75 segments** to **three annotators**.

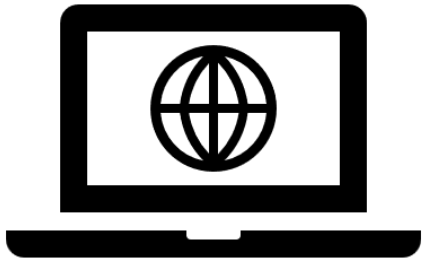
universität wien

Error Analysis (cont.)

Source (EN)	Reference (RO)	MT (RO)	Error category
Its utilisation is proposed within the anesthesiologist and/or intensivist scope of practice.	Utilizarea acestei metode este recomandată în practica anestezică și de terapie intensivă.	Utilizarea sa este propusă în domeniul de aplicare al **terapiei anestezice** și/sau **intensiviste**.	**Partial error** **Other error (Hallucination)**

universität wien

Error Analysis (cont.)

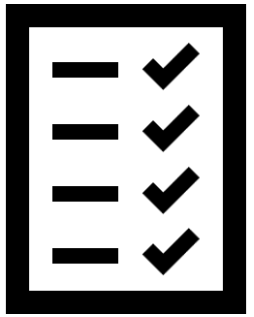


Setup

- Trados Studio 2021
- Qualitivity API plugin

Annotators (3)

- Collaborative annotation
- Native RO speakers
- 1 annotator with medical translation experience



Error typology

- Haque *et al*. (2019)
- Severity ranking: Minor < Major < Critical

universität wien

Error Analysis (cont.)

Terminology-related errors: **98** MBart, and **64** fine-tuned MBart.

Error type	MBart ↓	Fine-tuned MBart ↓
Partial error	41	**23**
Source term copied	22	**9**
Inflectional error	**2**	4
Reorder error	**1**	3
Disambiguation issue in target	14	**6**
Incorrect lexical selection	9	**6**
Term drop	0	0
Other error	**9**	13

universität wien

Error Analysis (cont.)

Terminology-related errors: **98** MBart, and **64** fine-tuned MBart.

Error type	MBart↓	Fine-tuned MBart↓
Partial error	41	**23**
Source term copied	22	**9**
Inflectional error	**2**	4
Reorder error	**1**	3
Disambiguation issue in target	14	**6**
Incorrect lexical selection	9	**6**
Term drop	0	0
Other error	**9**	13

universität wien

Error Analysis (cont.)

Other error category (fine-tuned MBart):

1. **Translation** of source term over **borrowing**
2. **Acronym recomposition** (borrowed and translated)
3. **Hallucinations**

universität wien

Error Analysis (cont.)

Other error category (fine-tuned MBart):

1. **Translation** of source term over **borrowing**

Source	Target (fine-tuned MBart)
"The global cortical connectivity increased during the **burst periods**."	"Conectivitatea corticală globală a crescut în timpul perioadelor de **arsură**." **[perioadelor de burst]**

universität wien

Error Analysis (cont.)

Other error category (fine-tuned MBart):

2. **Acronym recomposition** (borrowed and translated)

Source	Target (fine-tuned MBart)
"[...] of bioproduction of free radicals (**FR**) are significantly increasing in polytrauma patients."	"[...] bioproducţie a radicalilor liberi (**RF**) cresc semnificativ la pacienţii cu politrauma." **[RL]**

universität wien

Error Analysis (cont.)

Other error category (fine-tuned MBart):

3. Hallucinations

Source	Target (fine-tuned MBart)
"[...] the choice for femoral shaft stabilization by **intramedullary nailing** represents a safe option."	"[...] opţiunea stabilizării căilor femurale prin **nailing intramedullar** reprezintă o opţiune sigură." **[tijă centromedulară]**

universität wien

Error Analysis (cont.)

Other error category (MBart):

Translation over borrowing (1) and acronym recomposition (2),
but
no hallucinations

Source	Target (MBart)
(1) "[…] during the **burst** periods"	"[…] în timpul perioadelor de ardere" **[burst]**
(2) "[…] red blood cells units/48 h (**RBC**)"	"[…] unități de celule roșii/48 h (CCR)" **[RBC]**

universität wien

Error Analysis (cont.)

***Reorder error** category*:

1. **MBart:**
   - "[…] the DCO shock group" – "[…] grupul **de șoc DCO**" **[grupul DCO cu șoc]**
2. **Fine-tuned MBart:**
   - "[…] the DCO shock group" – "[…] grupul **cu șoc DCO**" **[grupul DCO cu șoc]**
   - "Similar results for: **ICU LOS**" – "Rezultate similare pentru: **LOS ICU**" **[ICU LOS]**

universität wien

Error Analysis (cont.)

***Inflectional error** category*:

1. **Both:**
   - "[…] anthropometric measurements […]" - "[…] **măsurarea antropometrică**" **[măsurătorile antropometrice]**
2. **MBart:**
   - "[…] the local field potentials" – "[…] potenţialele locale **ale câmpului**" **[de câmp]**
3. **Fine-tuned MBart:**
   - "[…] elective laparoscopic bariatric surgery" – "[…] intervenţie chirurgicală bariatrică laparoscopică **elective**" **[electivă]**
   - "duloxetine and venlafaxine" – "duloxetin**ă** şi venlafaxin**ă**" **[**duloxetin**a** şi venlafaxin**a]**

universität wien

Conclusions

- Impact of domain adaptation on MBart in the medical domain for English-Romanian
- Fine-tuned MBart outperforms the general model with **automatic metrics** and produces **fewer errors** related to **terminology**

universität wien

Future Work

- Collaborative annotation against **gold standard reference** texts.
  - Marking **over-** and **under-translation** in reference texts.
- Quantify the impact of fine-tuning on other error types present in the **Multidimensional Quality Metrics** Core (Lommel et al., 2013).
- Extend the **Byte-Pair Encoding** vocabulary in MBart to cope with in-domain terminology.

universität wien

References

- CLARIN:EL. (2015). EMEA Corpus. CLARIN:EL. http://hdl.handle.net/11500/ATHENA-0000-0000-25DB-0 Esperança-Rodier, E., Brunet-Manquat, F., & Eady, S. (2019, November).
- Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, F., Wattenberg, M., Corrado, G., Hughes, M., & Dean, J. (2017). **Google's Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation.** Transactions of the Association for Computational Linguistics.
- Haque, R., Hasanuzzaman, M., & Way, A. (2019). ***Investigating Terminology Translation in Statistical and Neural Machine Translation: A Case Study on English-to-Hindi and Hindi-to-English.*** Proceedings of the International Conference on Recent Advances in Natural Language Processing.
- Liu, Y., Gu, J., Goyal, N., Li, X., Edunov, S., Ghazvininejad, M., Lewis, M., & Zettlemoyer, L. (2020). ***Multilingual Denoising Pre-training for Neural Machine Translation.*** Transactions of the Association for Computational Linguistics
- Lommel, A. R., Burchardt, A., & Uszkoreit, H. (2013). ***Multidimensional quality metrics: A flexible system for assessing translation quality.***

Thank you and questions

HAITrans - Human and Artificial Intelligence in Translation

HAITrans - Human and Artificial Intelligence in Translation - is a research group based in the University of Vienna Centre for Translation Studies. It investigates the behavioural and cognitive effects which technologies such as machine translation and automatic speech recognition and synthesis have on translators, as well as their impact on the profession, practice, training and society at large.

At present, the core research areas of the Vienna HAITrans Group are:

1. Effects of speech technologies (Speech to Text and Text to Speech) on translation, revision and post-editing machine translation (PEMT) tasks (R1)
2. Technology-supported translation, revision, and PEMT practices (R2)
3. (Translation) technology for accessibility (R3)
4. Translation technology didactics (R4)

In our qualitative and quantitative investigations we use data gathered via eye-tracking, questionnaires, focus groups, corpora, and translation environment tool metrics. We also collaborate with academic partners, international organisations, language service providers, dedicated professional associations and cultural-sector partners.

Contact

Centre for Translation Studies (ZTW)
Kolingasse 14-16
1090 Vienna
T: +43-1-4277-58080
haitrans@univie.ac.at

Miguel Rios
miguel.angel.rios.gaona@univie.ac.at

Raluca Chereji
raluca-maria.chereji@univie.ac.at

Alina Secară
alina.secara@univie.ac.at

Dragoș Ciobanu
dragos.ioan.ciobanu@univie.ac.at

HAITrans research group
https://haitrans.univie.ac.at/