

LLMTrad-IBE: Large Language Models for translating Low-Resource Languages of the Iberian Peninsula^{*}

Antoni Oliver¹, Sergi Álvarez-Vidal²

¹Universitat Oberta de Catalunya (UOC), Rambla del Poblenou, 156 08018 Barcelona, Catalonia (Spain)

²Universitat Autònoma de Barcelona (UAB), Edifici K s/n, Plaça del Coneixement, 08193 Bellaterra, Catalonia (Spain)

Abstract

LLMTrad-IBE is a research initiative within the AI-TraLow coordinated project that focuses on developing and evaluating Large Language Models (LLMs) for the translation of low-resource Romance languages of the Iberian Peninsula, specifically Asturian, Aragonese, Aranese and Eonavian. The project focuses on specializing decoder-only architectures for these specific varieties to overcome the performance limitations of general-purpose models. Key research areas include the creation of high-quality evaluation benchmarks (FLORES+ and NTREX), the development of enhanced linguistic resources such as expanded Apertium dictionaries, and the implementation of knowledge distillation to ensure model efficiency. By comparing distilled decoder-only models against traditional encoder-decoder architectures, LLMTrad-IBE seeks to identify the most effective strategies for preserving the digital viability of low-resource Iberian languages. The project follows an Open Science paradigm, ensuring all models and tools are released under permissive licenses.

Keywords

Large Language Models, Machine Translation, Romance Languages

1. Project information

- Title: LLMTrad-IBE: Large Language Models for translating Low-Resource Languages of the Iberian Peninsula (subproject of AI-TraLow: AI-Driven Translation for Low-Resource Languages and Cultures).
- Founding institution: MCIU/ AEI/ 10.13039/ 501100011033/ FEDER, UE.
- Reference: PID2024-158157OB-C33
- Participating institutions: Universitat Oberta de Catalunya (coordinator of the subproject); Universidad de Oviedo; Universitat Autònoma de Barcelona; Universidad de Zaragoza
- Duration: 3 years (2025-2028)

2. Introduction

The LLMTrad-IBE project represents a strategic evolution in the development of machine translation (MT) systems for low-resource languages. This initiative is the direct successor to the TAN-IBE project (Neural Machine Translation for the Romance Languages of the Iberian Peninsula) [1], which successfully established a foundation for these languages using Neural Machine Translation (NMT) architectures.

While TAN-IBE focused on the capabilities of encoder-decoder neural models, the current technological landscape has been fundamentally reshaped by the emergence of Large Language Models (LLMs). LLMTrad-IBE marks a significant paradigm shift, transitioning from traditional NMT approaches to the exploration and adaptation of decoder-only LLM architectures. This transition is motivated by the

SEPLN 2026: 42nd International Conference of the Spanish Society for Natural Language Processing, León, Spain, 22-25 September 2026.

^{*} Subproject of AI-TraLow.

^{*} Corresponding author.

[†] These authors contributed equally.

✉ aoliverg@uoc.edu (A. Oliver); Sergi.Alvarez@uab.cat (S. Álvarez-Vidal)

🌐 <https://recerca.uoc.edu/investigadores/882326/detalle> (A. Oliver);

<https://portalrecerca.uab.cat/ca/persons/sergi-alvarez-vidal/> (S. Álvarez-Vidal)

🆔 0000-0001-8399-3770 (A. Oliver); 0000-0002-6355-4559 (S. Álvarez-Vidal)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

hypothesis that the few-shot learning capabilities and the vast linguistic knowledge encoded in LLMs can further bridge the gap for languages where parallel data is extremely scarce.

As a subproject within the AI-TraLow coordinated framework, LLMTrad-IBE focuses specifically on four Romance varieties: Asturian, Aragonese, Aranese and Eonavian. These languages face a critical digital divide due to their limited representation in the training datasets of general-purpose models. The project emphasizes a strong synergy with the other subprojects within AI-TraLow; while all three share similar overarching goals regarding low-resource machine translation, they strategically focus on different language families. Specifically, the University of Alicante (UA) leads the research on Mayan languages, and the Barcelona Supercomputing Center (BSC) focuses on languages of migrant communities in vulnerable situations. This coordinated structure allows for a cross-fertilization of methodologies and tools while ensuring specialized treatment for each linguistic group. Within this framework, LLMTrad-IBE aims to develop linguistic resources (such as monolingual, comparable and parallel corpora, and dictionaries), to fine-tune LLMs for translation for these languages.

2.1. Languages of the Project

The LLMTrad-IBE subproject focuses on four Romance varieties of the Iberian Peninsula that share a common trait: they are highly under-resourced in the digital sphere and face significant challenges regarding standardization and official recognition.

- **Asturian:** Spoken in the Principality of Asturias, it belongs to the Astur-Leonese linguistic group. It is estimated to have 100,000 native speakers and 450,000 second-language speakers. Although it lacks official status, it is protected under regional statutes and regulated by the Academy of the Asturian Language (ALLA). For MT purposes, the project focuses on the standard variety, a variety with a well-defined grammar and orthography but still suffering from a lack of large-scale parallel corpora.
- **Aragonese:** This language is the sole modern survivor of medieval Navarro-Aragonese that remains distinct from Spanish. It is currently spoken by approximately 12,000 people in the northern valleys of Aragon (Pyrenees). Aragonese is characterized by high dialectal fragmentation, but a significant added challenge for NLP is the historical diversity of competing orthographic standards. Although various systems have coexisted for decades, a new unified orthography was recently approved by the Academia Aragonesa de la Lengua. While this standard is expected to become the official and most widely used norm, it still coexists with previous ones, requiring LLMs to be robust enough to handle orthographic variation in existing digital texts. The project aims to consolidate resources for this variety, which has historically been relegated to oral contexts.
- **Aranese:** A standardized variety of Pyrenean Gascon (Occitan), Aranese is spoken in the Val d’Aran by a small community of approximately 5,000 to 7,000 habitual speakers (within a population of about 10,000 inhabitants). Despite this reduced number of speakers, its linguistic vitality and strategic importance are remarkable, as it is currently the only variety of the Occitan language—which has an estimated total of 500,000 to 600,000 speakers across southern France and parts of Italy—to enjoy full official status. Since 2010, it has been the third official language of Catalonia. This official recognition has fostered the creation of administrative, legal, and educational resources, providing a high-quality foundation for translation tasks that can potentially benefit the broader Occitan-speaking domain through cross-lingual transfer.
- **Eonavian:** Spoken in the westernmost area of Asturias, between the Eo and Navia rivers, Eonavian is a set of varieties traditionally classified within the Galician-Portuguese domain but exhibiting a strong transitional character toward Astur-Leonese. This *dialect continuum* status makes Eonavian a particularly interesting case for studying cross-lingual transfer in LLMs. Crucially, Eonavian is regulated by the Academy of the Asturian Language (ALLA), which has led notable efforts toward its standardization and the creation of formal linguistic and educational resources. Despite the political-linguistic debates regarding its classification, these institutional efforts provide a structured orthographic and grammatical framework that is essential for developing consistent training datasets and language identification tools.

3. Goals

The primary goal of the LLMTrad-IBE subproject is to advance the development of Machine Translation (MT) systems for low-resource Romance languages of the Iberian Peninsula by leveraging the capabilities of decoder-only Large Language Models (LLMs). This general objective is subdivided into the following specific objectives:

1. **Development of specialized linguistic resources:** To compile and curate high-quality monolingual, comparable, and parallel corpora for Asturian, Aragonese, Aranese and Eonavian, ensuring a solid foundation for model adaptation.
2. **Enlargement of lexical resources:** To automatically expand and refine existing Apertium dictionaries (both monolingual and transfer) to compensate for the scarcity of parallel data in these language pairs.
3. **LLM adaptation for low-resource Romance languages:** To research and implement fine-tuning strategies and in-context learning techniques that incorporate explicit linguistic information into LLMs to improve translation accuracy.
4. **Knowledge distillation and efficiency:** To develop compact and efficient translation models through knowledge distillation, enabling the deployment of high-quality systems in standard computing environments.
5. **Comprehensive evaluation and benchmarking:** To create and release new evaluation benchmarks (based on FLORES+ and NTREX) and conduct a rigorous comparison between distilled decoder-only models and traditional encoder-decoder architectures.
6. **Promotion of Digital Inclusion:** To ensure the digital survival of these low resource languages by releasing all resulting models, tools, and datasets under open-source licenses.

4. Methodology and work packages

The project’s methodology is based on a three-phase iterative cycle that will be executed twice over the three-year duration. This iterative approach allows for the continuous refinement of models as new data and techniques become available. These phases include: (i) data compilation and curation, (ii) research into LLM training and fine-tuning, and (iii) the release and evaluation of high-performance models.

4.1. Iterative Development Cycle

In the first phase, the focus is on the acquisition of high-quality resources, ranging from document scanning and OCR to targeted web crawling. A key innovation in this stage, implemented in a different subproject, is the development of automated mechanisms to detect and filter machine-generated or synthetically translated content, ensuring the integrity of the training sets.

The second phase is dedicated to the core research of LLM adaptation. We explore both in-context learning (ICL) and fine-tuning strategies, specifically investigating how to integrate symbolic knowledge—such as dictionaries and grammatical rules—into neural architectures. Finally, the third phase involves the human and automatic evaluation of the resulting systems to ensure they meet the quality standards required for public release.

4.2. Work Packages and Subproject Tasks

The subproject LLMTrad-IBE is integrated into the global work plan through the following Work Packages¹:

¹Please, note that the numbering refers to the WP in the coordinated project, and not all of them apply to our subproject.

4.2.1. WP1. Data Acquisition and Curation

In this WP we will the hybridization of rule-based and neural approaches, focusing on exploiting underutilized resources such as components from existing Apertium rule-based MT systems. Apertium has demonstrated to be very useful for the creation of synthetic and parallel corpora [2]. To that end, we plan to automatically enlarge the monolingual and transfer dictionaries of Apertium using existing lexical resources and monolingual corpora. The generated corpora will be used for the fine-tuning of our Large Language Models (LLMs) for translation. Furthermore, we will develop methods to integrate monolingual dictionaries and grammar books directly into LLM prompts and fine-tuning processes.

Data acquisition will involve the crawling and language identification of documents from the Internet. We will collect web-based documents and create algorithms for efficient crawling; since many of the targeted languages lack existing tools for automatic language identification, we will develop specialized language identification tools as required to address this gap.

Regarding evaluation, we plan to create new versions of the FLORES+ [3] evaluation datasets for Aragonese, Aranese, and Eonavian, by asking professional translators to translate the English version from scratch. This approach aims to eliminate the biases introduced by relying on machine translation, ensuring higher quality and linguistic accuracy. Similarly, we aim to create the NTREX dataset² [4] for the same languages, as well as for Asturian. While the aforementioned datasets are segment-level, LLMs are capable of effectively handling larger contexts; to address this, we plan to develop evaluation datasets at the paragraph level for the same languages. These outcomes will support the integration of these languages into multilingual experiments, significantly advancing research on LLMs for translation both into and from these languages.

4.2.2. WP2. Leveraging Linguistic Information in LLMs

This Work Package explores the hybridization of rule-based and neural approaches to overcome the limitations of data-scarce scenarios. We focus on exploiting underutilized resources, such as components from existing Apertium rule-based MT systems improved in WP1, and developing methods to integrate both structured and unstructured linguistic information directly into LLM workflows.

Regarding structured linguistic information, we focus on the systematic integration of bilingual dictionaries and transfer rules. These resources provide precise lexical and morphological guidance that LLMs often lack for low-resource languages. Furthermore, this structured data serves as a foundation for the generation of high-quality synthetic data; by leveraging rule-based engines to produce diverse and grammatically correct sentence pairs, we can significantly augment the available training and fine-tuning sets.

In parallel, we investigate the use of unstructured linguistic information, such as descriptive grammar books and language learning materials. Unlike rigid dictionaries, these resources contain rich, prose-based explanations of syntax and usage. We develop advanced prompting techniques and fine-tuning strategies to "teach" the model these linguistic constraints, effectively providing the LLM with a pedagogical context that enhances its translation reasoning. By combining these diverse knowledge sources, WP2 aims to create a more robust and linguistically grounded translation framework for the Romance languages of the Iberian Peninsula.

4.2.3. WP4. Strategies for Training and Adapting LLMs

This Work Package focuses on developing scalable, resource-efficient, and ethically robust solutions for low-resource machine translation. A key priority is the exploration of knowledge distillation, where the capabilities of massive LLMs are transferred into more compact and efficient architectures—such as distilled decoder-only or specialized encoder-decoder models—capable of running on standard hardware. To complement this, the WP is structured around the following specialized tasks:

²<https://github.com/MicrosoftTranslator/NTREX>

- **Strategies for enriching LLMs with synthetic corpora.** In order to augment the training data for data-scarce languages, we plan to generate synthetic parallel corpora through backtranslation, that is, obtaining synthetic parallel data where the target part is of good quality, by using an initial MT engine to translate monolingual text in the low-resource language into the target language. The initial systems used to create these synthetic datasets may be small encoder-decoder engines or existing legacy rule-based MT systems, such as Apertium. Currently, Apertium systems are available for: Spanish $\rightarrow$ Asturian, Aragonese $\leftrightarrow$ Spanish, Aragonese $\leftrightarrow$ Catalan, Aranese $\leftrightarrow$ Spanish, and Aranese $\leftrightarrow$ Catalan. For the Asturian $\rightarrow$ Spanish direction, where Apertium is unavailable, we will use the best-performing neural MT system developed in the *TAN-IBE* project. Whenever possible, back-translated corpora will be utilized [5, 6, 7], alongside the compiled parallel corpora.
- **Detecting and mitigating gender bias in MT with LLMs.** Ensuring fairness is a core objective. We will use benchmarks and datasets, such as WinoMT [8], containing sentences with ambiguous or explicit gender contexts to assess bias in translations. To identify both explicit and subtle gender biases and to mitigate them, we will incorporate techniques from [9] and explore a variety of approaches.
- **Evaluation of LLMs fine-tuning strategies.** Several strategies for fine-tuning LLMs have emerged in recent years, including parameter-efficient fine-tuning [10, 11, 12], instruction tuning [13], and representation fine-tuning [14]. In this task, we aim to evaluate all these strategies, as well as any new techniques that may emerge, to determine which is most suitable for adapting LLMs to the specific linguistic constraints of low-resource languages. The evaluation will rely on standard automatic metrics (such as BLEU, chrF, and COMET [15]) to ensure a rigorous comparison of performance across different fine-tuning paradigms.

4.2.4. WP5. Building Models for the Iberian Romance Languages

This is the core application package for the subproject, where the methodologies and linguistic resources developed in previous WPs are applied to create the final translation models. In this WP, we will develop Large Language Models (LLMs) for translating the targeted Iberian Romance languages, applying the lessons learnt regarding data curation and hybrid architectures. The work is structured to determine whether fine-tuning encoder-decoder systems or specialized LLMs is more effective for these specific linguistic varieties, guiding future resource allocation. The main tasks of this WP are:

- **Training and release of LLMs for low-resource languages of the Iberian Peninsula.** In this task, we will build MT systems based on LLMs for Asturian, Aragonese, Aranese, and Eonavian, utilizing the curated resources and synthetic corpora previously developed. This task will be executed multiple times throughout the project, with each iteration showcasing the advancements and technologies developed up to that point, ensuring a continuous release of improved models to the community.
- **Qualitative evaluation of LLM translations.** We plan to automatically evaluate all the MT systems developed in the project by focusing on the specific characteristics of the translations they produce. Specifically, we will assess lexical diversity using metrics such as TTR and MTLTD [16], literalness based on monotonicity and word alignment [17], and the degree of interference from the source language by using perplexity on language models trained on part-of-speech sequences [18].
- **Evaluation and comparison of encoder-decoder models and LLMs for translation.** Human evaluation is crucial to understand error typologies, despite its complexity and potential for low inter-annotator agreement [19]. In this task, we will conduct a rigorous comparison between distilled decoder-only models and traditional architectures. We will follow an evaluation methodology grounded in explicit error analysis [20], based on the Multidimensional Quality Metrics (MQM) framework [21]. This qualitative approach will allow us to identify the strengths and weaknesses of LLMs versus traditional NMT in the context of the Iberian Romance languages.

5. Project Deliverables and Resources

The LLMTrad-IBE subproject is committed to the generation of tangible and open-access resources. Following the work plan, the following deliverables will be developed and released:

- **Corpora and Lexical Resources (WP1 & WP2):**

- Specialized language identification tools for the targeted Iberian varieties.
- Automatically expanded monolingual and transfer dictionaries for the Apertium platform.
- New versions of the *FLORES+* and *NTREX* evaluation datasets for Aragonese, Aranese, and Eonavian, translated from scratch by professional translators.
- Evaluation datasets at the paragraph level to assess long-context handling in LLMs.
- High-quality parallel and synthetic corpora generated through back-translation and rule-based hybridization.

- **Models and Optimization (WP4 & WP5):**

- Specialized Large Language Models (LLMs) fine-tuned for Asturian, Aragonese, Aranese, and Eonavian.
- Distilled, compact decoder-only models and specialized encoder-decoder architectures optimized for standard computing environments.
- Documentation on the effectiveness of different fine-tuning strategies (PEFT, Instruction Tuning, ReFT) for low-resource Romance languages.

- **Evaluation and Quality Metrics (WP4 & WP5):**

- Benchmarks for detecting and mitigating gender bias in translation (based on *WinoMT*).
- A comprehensive qualitative evaluation framework based on MQM (Multidimensional Quality Metrics) for error typology analysis.
- Scripts and tools for the automatic assessment of translation quality, including lexical diversity (TTR, MTLT), literalness, and source language interference.

All the aforementioned resources will be made available under permissive open-source licenses through the project’s official GitHub repositories and the *LLMTrad-IBE* portal:

- Github repository of the AITraLow project³
- Github repository of the LLMTrad-IBE project⁴

6. Conclusions

The LLMTrad-IBE project establishes a comprehensive roadmap for integrating low-resource Romance languages into the current era of Large Language Models. By transitioning from the foundations laid by the TAN-IBE project to state-of-the-art decoder-only architectures, this initiative addresses the critical digital gap that threatens the linguistic diversity of the Iberian Peninsula.

Our approach combines the generative power of LLMs with the structural precision of rule-based systems like Apertium, demonstrating that hybridization is the most viable path forward in data-scarce scenarios. The development of high-quality, human-translated benchmarks such as *FLORES+* and *NTREX* from scratch is essential to ensure that progress in machine translation is measured with accuracy and without the inherent biases of existing automated systems. Furthermore, the exploration of knowledge distillation ensures that the resulting technology remains accessible and sustainable, allowing low-resource language communities to deploy efficient models on standard hardware.

Ultimately, LLMTrad-IBE is firmly committed to the principles of Open Science. By releasing all models, curated corpora, and expanded lexical resources under permissive licenses, we aim to provide the scientific community and local linguistic institutions with the necessary tools to ensure that Asturian, Aragonese, Aranese and Eonavian are not only preserved but also digitally competitive in the global landscape of artificial intelligence.

The project has a duration of 3 years and has started on September 2025.

³<https://transducens.github.io/ai-tralow/>

⁴<https://llmtrad-ibe.github.io/>

Acknowledgments

This work is supported by the project granted by MCIU/AEI/10.13039/501100011033/FEDER,UE with reference PID2024-18157OB-C33.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

- [1] A. Oliver, M. Vázquez, M. Coll-Florit, S. Álvarez, V. Suárez, C. Aventín-Boya, C. Valdés, M. Font, A. Pardos, Tan-ibe: Neural machine translation for the romance languages of the iberian peninsula, in: Proceedings of the 24th Annual Conference of the European Association for Machine Translation, 2023, pp. 495–496.
- [2] F. Sánchez-Martínez, J. A. Pérez-Ortiz, A. Galiano-Jiménez, A. Oliver, Findings of the wmt 2024 shared task translation into low-resource languages of spain: Blending rule-based and neural systems, in: Proceedings of the Ninth Conference on Machine Translation, 2024, pp. 684–698.
- [3] J. A. Pérez-Ortiz, F. Sánchez-Martínez, V. M. Sánchez-Cartagena, M. Esplà-Gomis, A. Galiano-Jiménez, A. Oliver, C. Aventín-Boya, A. Pardos, C. Valdés, J. L. S. Socasau, et al., Expanding the flores+ multilingual benchmark with translations for aragonese, aranese, asturian, and valencian, in: Proc. of the 9th Conference on Machine Translation, 2024, pp. 547–555.
- [4] C. Federmann, T. Kocmi, Y. Xin, Ntrex-128–news test references for mt evaluation of 128 languages, in: Proc. of the First Workshop on Scaling Up Multilingual Evaluation, 2022, pp. 21–24.
- [5] R. Sennrich, B. Haddow, A. Birch, Improving neural machine translation models with monolingual data, in: Proc. of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Berlin, Germany, 2016, pp. 86–96. URL: <https://aclanthology.org/P16-1009/>. doi:10.18653/v1/P16-1009.
- [6] S. Edunov, M. Ott, M. Auli, D. Grangier, Understanding back-translation at scale, in: Proc. of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, 2018, pp. 489–500. URL: <https://www.aclweb.org/anthology/D18-1045>. doi:10.18653/v1/D18-1045.
- [7] M. Graça, Y. Kim, J. Schamper, S. Khadivi, H. Ney, Generalizing back-translation in neural machine translation, in: Proc. of the Fourth Conference on Machine Translation (Volume 1: Research Papers), Florence, Italy, 2019, pp. 45–52. URL: <https://aclanthology.org/W19-5205/>. doi:10.18653/v1/W19-5205.
- [8] G. Stanovsky, N. A. Smith, L. Zettlemoyer, Evaluating gender bias in machine translation, in: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL), Association for Computational Linguistics, Florence, Italy, 2019, pp. 1679–1684. URL: <https://aclanthology.org/P19-1164/>. doi:10.18653/v1/P19-1164.
- [9] H. Gonen, K. Webster, Automatically identifying gender issues in machine translation using perturbations, in: Findings of the Association for Computational Linguistics: EMNLP 2020, 2020, pp. 1991–1995.
- [10] N. Houlsby, A. Giurgiu, S. Jastrzebski, B. Morrone, Q. De Laroussilhe, A. Gesmundo, M. Attariyan, S. Gelly, Parameter-efficient transfer learning for NLP, in: International Conference on Machine Learning, 2019, pp. 2790–2799.
- [11] E. J. Hu, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen, et al., Lora: Low-rank adaptation of large language models, in: International Conference on Learning Representations, 2021.
- [12] Y. Xu, L. Xie, X. Gu, X. Chen, H. Chang, H. Zhang, Z. Chen, X. Zhang, Q. Tian, QA-LoRA: Quantization-aware low-rank adaptation of large language models, 2023. [arXiv:2309.14717](https://arxiv.org/abs/2309.14717).
- [13] S. Zhang, L. Dong, X. Li, S. Zhang, X. Sun, S. Wang, J. Li, R. Hu, T. Zhang, F. Wu, et al., Instruction tuning for large language models: A survey, 2023. [arXiv:2308.10792](https://arxiv.org/abs/2308.10792).

- [14] Z. Wu, A. Arora, Z. Wang, A. Geiger, D. Jurafsky, C. D. Manning, C. Potts, ReFT: Representation finetuning for language models, 2024. [arXiv: 2404.03592](https://arxiv.org/abs/2404.03592).
- [15] R. Rei, C. Stewart, A. C. Farinha, A. Lavie, Comet: A neural framework for mt evaluation, 2020. URL: <https://arxiv.org/abs/2009.09025>. [arXiv: 2009.09025](https://arxiv.org/abs/2009.09025).
- [16] E. Vanmassenhove, D. Shterionov, A. Way, Lost in translation: Loss and decay of linguistic richness in machine translation, in: Proc. of Machine Translation Summit XVII: Research Track, 2019, pp. 222–232.
- [17] V. Raunak, A. Menezes, M. Post, H. H. Awadalla, Do gpts produce less literal translations?, in: Proc. of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), 2023, pp. 1041–1050.
- [18] A. Toral, Post-edits: an exacerbated translationese, in: Proc. of Machine Translation Summit XVII: Research Track, Dublin, Ireland, 2019, pp. 273–281. URL: <https://aclanthology.org/W19-6627/>.
- [19] M. Snover, B. Dorr, R. Schwartz, L. Micciulla, J. Makhoul, A Study of Translation Edit Rate with Targeted Human Annotation, 2006. URL: https://www.cs.umd.edu/~snover/pub/amta06/ter_amta.pdf<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.129.4369&rep=rep1&type=pdf>. doi:10.1.1.129.4369.
- [20] M. Freitag, G. Foster, D. Grangier, V. Ratnakar, Q. Tan, W. Macherey, Experts, errors, and context: A large-scale study of human evaluation for machine translation, Transactions of the Association for Computational Linguistics 9 (2021) 1460–1474. URL: <https://aclanthology.org/2021.tacl-1.87/>. doi:10.1162/tacl_a_00437.
- [21] A. Burchardt, Multidimensional quality metrics: a flexible system for assessing translation quality, in: Proc. of Translating and the Computer 35, 2013, pp. 28–29. URL: <https://aclanthology.org/2013.tc-1.6/>.