# FROM STANDARDS TO HYBRID WORKFLOWS: AUTOMATIC POST-EDITING AND TRANSLATION QUALITY ASSESSMENT FOR ENGLISH-ROMANIAN MEDICAL TEXTS

**Zoltan KOVACS**
West University of Timișoara, Romania
**Daniel DEJICA**
Politehnica University Timișoara, Romania

**Abstract:** Medical translation in high-risk domains such as healthcare requires workflows that combine efficiency with rigorous quality assurance. This article proposes a hybrid model integrating Automatic Post-Editing (APE) and Translation Quality Assessment (TQA) to improve English-Romanian medical translations. Building on ISO 18587 and error taxonomies such as Multidimensional Quality Metrics (MQM) and Dynamic Quality Framework (DQF), the framework introduces APE as an intermediate error-reduction layer and hybrid TQA combining human annotation with automated quality estimation. The study proposes an eight-stage workflow aligned with international standards and tailored to risk-sensitive genres like patient information leaflets and discharge summaries. In addition, a theoretical pilot protocol is presented to validate the workflow and evaluate accuracy, fluency, and efficiency across three translation pipelines: human translation, Neural Machine Translation (NMT) with post-editing, and NMT with APE plus post-editing. By bridging academic research and professional practice, this contribution advances a standards-driven, implementable approach for technology-mediated medical translation, addressing Romanian-specific gaps in resources, training, and workflow design.

**Keywords:** medical translation; automatic post-editing; ISO 18587; translation quality assessment; neural machine translation.

## 1. Introduction

Neural Machine Translation (NMT) has transformed professional translation workflows, offering unprecedented speed and fluency. However, in high-stakes domains such as medicine, where errors in terminology or dosage instructions can compromise patient safety, rigorous quality assurance remains indispensable (Rojas Plata & Castro Sánchez, 2024). Recent systematic reviews confirm these concerns, highlighting variability in NMT performance across clinical contexts and the need for human oversight (Karakus et al., 2025). For English-Romanian medical translation, these challenges are amplified by the scarcity of domain-specific resources and the complexity of multilingual healthcare communication.

Post-editing (PE) and Translation Quality Assessment (TQA) are widely recognized as essential safeguards in technology-mediated translation (ISO 18587:2017; Nitzke & Hansen-Schirra, 2021). Yet, traditional workflows often struggle to balance efficiency with reliability. Recent advances in Automatic Post-Editing (APE) and hybrid evaluation models—combining human annotation with automated quality estimation—offer promising solutions (do Carmo et al., 2021; Specia, Scarton & Paetzold, 2018). These innovations align with international standards such as ISO 18587 and error taxonomies

like MQM and DQF, but their integration into Romanian practice remains limited (Kovacs & Dejica, 2025).

This article addresses that gap by proposing a hybrid workflow that integrates APE as an error-reduction layer within a structured pipeline, complemented by hybrid TQA combining human and semi-automated evaluation. The approach is tailored to risk-sensitive genres such as patient information leaflets and discharge summaries, where accuracy and clarity are critical. Specifically, the study aims to: (1) synthesize current standards and models relevant to PE, APE, and TQA in medical translation; (2) outline an implementable APE–TQA workflow aligned with ISO 18587 and risk-based acceptance criteria; and (3) present a pilot protocol for evaluating the proposed workflow in terms of accuracy, fluency, and efficiency.

Unlike previous studies that focus on descriptive analysis, this paper operationalizes a standards-aligned workflow for empirical validation. By moving beyond descriptive analysis toward theory-building in translation process design, this contribution bridges academic research and professional practice. It offers principles for balancing automation and human oversight, ensuring that technological innovation supports—not compromises—the ethical and communicative demands of medical translation.

## 2. Background

The integration of machine translation (MT) technologies into professional workflows has transformed the landscape of specialized translation, yet it has also introduced new challenges for quality assurance. In high-stakes domains such as medical communication, where errors can have critical consequences, ensuring accuracy and reliability requires structured methodologies that combine automation with human oversight. This section provides an overview of three key components underpinning technology-mediated translation quality: post-editing standards and levels, which define the scope and requirements for refining MT output; automatic post-editing (APE), an emerging solution for reducing systematic errors through machine learning; and translation quality assessment (TQA) models, which offer frameworks for evaluating linguistic adequacy and functional equivalence. Together, these elements form the theoretical foundation for the hybrid workflow proposed in this study, aligning international standards with practical strategies for risk-sensitive genres such as English-Romanian medical texts.

### 2.1 Post-editing standards and levels

Post-editing (PE) has become a cornerstone of workflows involving machine translation (MT), particularly in high-risk domains such as medicine. The international standard ISO 18587:2017 defines requirements for post-editing MT output, distinguishing between light and full post-editing (ISO 18587:2017, 2017). Light PE ensures comprehensibility and factual accuracy, while full PE aims for quality comparable to human translation, addressing grammar, style, and formatting. Industry frameworks such as the TAUS Dynamic Quality Framework (DQF) and Multidimensional Quality Metrics (MQM) operationalize error categories and severity levels, enabling systematic evaluation and reporting (TAUS, n.d.; Lommel et al., 2013). Competence specifications for post-editors emphasize linguistic expertise, domain knowledge, and familiarity with MT systems (Nitzke & Hansen-Schirra, 2021).

### 2.2 Automatic Post-Editing (APE)

Automatic Post-Editing (APE) emerged as a response to recurring error patterns in MT output. Early approaches relied on rule-based and statistical methods to correct systematic errors (Lagarda et al., 2009). Recent developments have shifted toward neural architectures trained on triplets of source text, MT output, and human post-edits (do Carmo et al., 2021). APE systems can reduce repetitive errors, improve fluency, and support domain adaptation, particularly when integrated into workflows for specialized fields (Specia, Scarton & Paetzold, 2018). However, their effectiveness depends on factors such as segment length, error density, and the availability of high-quality training data—conditions that remain challenging for low-resource language pairs like English-Romanian.

### 2.3 Translation Quality Assessment (TQA) models

Quality assessment in translation has evolved from human-centric approaches to hybrid and automated models. Traditional frameworks such as House's functional-pragmatic model (House, 2015), Waddington's error-based approach (Waddington, 2003), and the American Translators Association (ATA) grading system (ATA, n.d.) emphasize communicative function, error severity, and transparency. Automated metrics such as BLEU (Papineni et al., 2002), METEOR (Banerjee & Lavie, 2005), and TER offer scalability but often fail to capture nuanced linguistic and functional adequacy. Hybrid models combining human annotation aligned with MQM/DQF and automated quality estimation (Specia, Scarton & Paetzold, 2018; Zhu, 2023) provide a balanced solution, enabling granular error analysis while leveraging computational tools for efficiency—particularly relevant for medical translation where accuracy and clarity are paramount.

## 3. Romanian context: gaps and opportunities

Medical translation in Romania occupies a critical yet underexplored niche within professional and academic discourse. Despite the global momentum in integrating machine translation (MT), post-editing (PE), and translation quality assessment (TQA) into workflows, Romanian practice remains fragmented and underdeveloped (Dejica et al., 2022; Kovacs & Dejica, 2025). Several structural and operational gaps persist: the scarcity of domain-specific bilingual corpora, limited adoption of ISO 18587 standards, and insufficient exposure to MQM/DQF-based error taxonomies in translator training programs.

Existing research confirms these deficits. Kovacs & Dejica (2025) mapped international standards and TQA models against Romanian practice, revealing minimal integration of post-editing protocols and quality frameworks in both academia and industry. Similarly, Crăineanu & Dejica (2025) emphasize the pedagogical gap, noting that English for Medical Purposes (EMP) curricula in Romanian universities rarely incorporate MT/PE activities or genre-based translation tasks. Curriculum development remains a critical priority, as recent studies advocate for integrating MT and post-editing competencies into translator education to meet industry demands (Grigoraș & Dejica, 2025b). This lack of structured training leaves practitioners ill-equipped to manage risk-sensitive genres such as patient information leaflets and discharge summaries.

Professional practice reflects similar shortcomings. While some language service providers experiment with MT tools, workflows often lack systematic quality assurance,

resulting in ad hoc post-editing and inconsistent error handling (Medical Language Service, 2025). The absence of Romanian-language medical corpora and terminological databases further limits the effectiveness of neural MT and Automatic Post-Editing (APE) systems, which rely on large, domain-specific datasets for optimal performance (Karakus et al., 2025).

Opportunities for improvement are significant. First, developing specialized corpora and terminological resources would enable domain adaptation of MT and APE systems, improving accuracy and reducing cognitive load during post-editing. Second, embedding ISO 18587 principles and MQM/DQF-based evaluation criteria into translator training curricula could enhance competence in high-stakes domains (ISO 18587:2017; Lommel et al., 2013). Third, collaborative initiatives between universities, healthcare institutions, and language service providers could support pilot projects and shared evaluation protocols, fostering a culture of evidence-based practice (Grigoraș & Dejica, 2025a; Mali & Dejica, 2025a; Mali & Dejica, 2025b).

Addressing these gaps requires a coordinated strategy that aligns international standards with local needs. By integrating structured post-editing workflows, hybrid TQA models, and targeted training programs, Romanian translation studies can position itself within the global discourse on technology-mediated translation quality while responding to the ethical and communicative demands of medical translation (Dejica et al., 2022; Kovacs & Dejica, 2025).

## 4. Proposed hybrid APE–TQA workflow

Translation workflows represent structured sequences of tasks designed to ensure efficiency, consistency, and quality in translation projects. They typically encompass three main stages: pre-production, production, and post-production, each involving specific activities such as project analysis, resource preparation, translation, revision, and final delivery (Pașcalău & Dejica, 2021). An efficient workflow is not only a technical necessity but also a strategic asset, enabling translation providers to optimize time, cost, and human resources while maintaining compliance with standards and client requirements (Dejica, 2016a; Dejica, 2016b; DG Translation, 2016). Research emphasizes that workflows should be adaptable to project size and complexity, integrating tools such as translation memories, terminology databases, and quality assurance protocols to support scalability and sustainability (Svoboda, Biel, & Łoboda, 2017; EAMT, 1998).

In this context, the proposed workflow builds on principles outlined in previous research on technology-driven translation processes, which emphasize the role of digital tools and automation in enhancing efficiency and quality (Dejica, Eugeni & Dejica- Cartiș, 2020). It uses established principles of translation quality assurance, aligning automation with human oversight to address the stringent requirements of high-risk domains such as medical translation. Designed in accordance with ISO 18587 standards, the workflow responds to the dual challenge of ensuring reliability in English–Romanian medical texts while mitigating the limitations of low-resource language pairs. It operationalizes a structured, eight-stage process that integrates machine translation and automatic post-editing with human intervention and hybrid quality assessment, thereby creating a scalable and standards-aligned model for improving accuracy, fluency, and efficiency in specialized translation workflows.

1. ***Pre-assessment***
   Texts are classified according to genre and risk level. Patient information leaflets and discharge summaries, for example, require full post-editing due to their critical nature, whereas less sensitive materials may allow light post-editing.
2. ***Resource Preparation***
   Essential resources include bilingual glossaries, domain-specific translation memories, and style guides. A Do-Not-Translate list and an MQM/DQF error profile are also established to ensure consistency and facilitate quality assessment.
3. ***Machine Translation and Automatic Post-Editing***
   Neural MT systems generate initial translations, which are then processed through APE models trained to correct systematic errors such as terminology inconsistencies, diacritic omissions, and punctuation issues. This step reduces repetitive errors before human intervention.
4. ***Human Post-Editing***
   Professional post-editors apply light or full post-editing according to the risk classification. ISO 18587 guidelines inform acceptance criteria, ensuring that factual accuracy, linguistic fluency, and formatting standards are met.
5. ***Hybrid Translation Quality Assessment***
   Quality is evaluated through a combination of human annotation and automated metrics. Human reviewers apply MQM/DQF categories to identify and classify errors, while automated quality estimation tools provide severity triage and predictive scoring.
6. ***Final Quality Assurance***
   Formatting, locale-specific conventions, and compliance checks are performed. This stage ensures that the final output meets both linguistic and regulatory requirements.
7. ***Metrics and Reporting***
   Performance indicators such as Translation Edit Rate (TER), MQM error rates, turnaround time, and cost differentials are documented. These metrics support continuous improvement and provide empirical evidence for workflow optimization.
8. ***Decision: Accept or Revise/Reject***
   Based on hybrid TQA results, the translation is either accepted for delivery or returned for revision. This decision point ensures compliance with quality standards and mitigates risks in high-stakes medical communication.

The workflow proposed in this study is illustrated through two complementary representations. Figure 1 provides a schematic overview of the eight stages involved in the hybrid APE–TQA process, highlighting the sequence of tasks from pre-assessment to metrics and reporting. Figure 2 offers a simplified graphical summary, designed for quick visualization of the core logic behind the workflow—how machine translation, automatic post-editing, and human post-editing interact with hybrid quality assessment to ensure compliance with standards. Together, these figures present both a detailed and an at-a-glance perspective on the proposed model, supporting its theoretical and practical applicability.

| Stage | Description |
|---|---|
| Pre-assessment | Genre selection and risk classification |
| Resource preparation | Glossaries, TM, style guide, MQM/DQF profile |
| MT + APE | Neural MT output corrected by APE models |
| Human post-editing | Light/full PE based on ISO 18587 |
| Hybrid TQA | Human MQM/DQF + automated QE |
| Final QA | Formatting, compliance, locale checks |
| Metrics & reporting | TER, MQM error rates, time, cost |
| Decision | Accept or revise/reject |

**Figure 1.** Hybrid APE–TQA Workflow. Schematic representation.

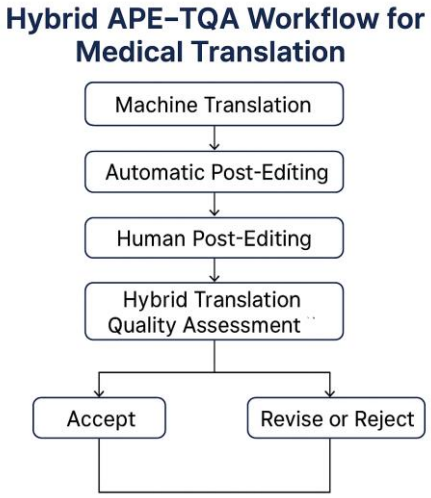


**Figure 2.** Hybrid APE–TQA Workflow. Simplified graphic summary.

## 5. Pilot protocol

To evaluate the effectiveness of the proposed hybrid APE–TQA workflow, a pilot study should be designed using a controlled experimental setting. The protocol is theoretical and intended for future applications to validate the hybrid workflow. It is recommended that the design adopt a within-subjects approach, enabling direct comparison of different translation pipelines on the same set of texts. Two medical genres should be selected for their practical relevance and linguistic complexity: patient information leaflets (PILs) and hospital discharge summaries. These genres combine specialized terminology with communicative constraints, making them suitable for assessing both accuracy and readability.

Three translation pipelines (Table 1) are recommended for testing:

- Pipeline A: Human translation followed by revision.
- Pipeline B: Neural Machine Translation (NMT) followed by human post-editing.

- Pipeline C: NMT combined with Automatic Post-Editing (APE) and subsequent human post-editing.

Each pipeline should be applied to approximately 20 documents per genre to ensure sufficient data for statistical analysis. The evaluation should employ Multidimensional Quality Metrics (MQM) for error categorization and severity scoring, complemented by Translation Edit Rate (TER) and Human Translation Edit Rate (HTER) for quantitative comparison. Additional measures should include time-on-task for post-editing and inter-rater agreement among reviewers.

The analysis should combine quantitative metrics (error counts, severity scores, effort indicators) with qualitative insights derived from error typology and reviewer feedback. Statistical tests such as paired comparisons are recommended to determine whether the hybrid workflow offers significant improvements in accuracy and efficiency compared to traditional approaches. Ethical considerations include anonymization of medical texts and the involvement of qualified reviewers with expertise in both translation and medical terminology.

| Pipeline | Description | Measures | Expected Outcome |
|---|---|---|---|
| A | Human translation + revision | MQM error rates, HTER, time-on-task | Baseline for quality and effort |
| B | NMT + human post-editing | MQM error rates, HTER, time-on-task | Reduced turnaround time vs. A |
| C | NMT + APE + human post-editing | MQM error rates, HTER, time-on-task | Further efficiency gains; error reduction in repetitive patterns |

**Table 1.** Pilot Protocol Overview

The pilot study should remain theoretical at this stage and is intended for future implementation to validate the proposed hybrid workflow. It is recommended that the design emphasize comparability across conditions and include both quantitative and qualitative dimensions. Quantitatively, paired statistical tests should be applied to error counts and measures of post-editing effort, enabling the identification of significant differences between workflows. Qualitatively, a detailed error typology should be compiled to capture patterns in terminology, syntax, and functional adequacy, offering insights into the nature and distribution of residual errors after automatic and human intervention.

## 6. Discussion and limitations

The integration of Automatic Post-Editing (APE) with hybrid Translation Quality Assessment (TQA) should be considered a recommended approach for improving English-Romanian medical translation workflows. Introducing APE as an intermediate layer is expected to mitigate repetitive and systematic errors—such as incorrect diacritics, punctuation inconsistencies, and capitalization—before human intervention. This reduction in error density should allow post-editors to concentrate on higher-level issues, including terminology accuracy and stylistic conformity. Furthermore, the hybrid TQA model, which combines human annotation with automated quality estimation, is recommended for ensuring reliability while maintaining scalability. Such an approach aligns with international standards and should respond effectively to the growing demand for cost-efficient yet safe workflows in medical communication.

However, several limitations must be acknowledged. The effectiveness of APE depends on the availability of high-quality training data, which remains scarce for Romanian medical texts; this limitation should be addressed in future research. While hybrid TQA enhances diagnostic precision, it introduces complexity in resource allocation and requires specialized expertise for error annotation and interpretation. Ethical considerations must also be observed, particularly regarding patient safety and confidentiality. Any use of clinical texts beyond publicly available or anonymized sources should comply strictly with data protection regulations and institutional review protocols. Finally, the proposed workflow should not be viewed as a replacement for qualified human reviewers; rather, it is recommended as a complementary mechanism that underscores the need for continuous training and professional development.

It should also be emphasized that the pilot protocol presented in this paper is theoretical and designed for future application. Its purpose is to provide a structured basis for empirical validation of the hybrid workflow, ensuring that subsequent implementations can be tested under controlled conditions before adoption in professional practice.

## 7. Conclusion

The study advances the discussion on machine-assisted medical translation by proposing a hybrid workflow that combines Automatic Post-Editing (APE) with a structured Translation Quality Assessment (TQA) approach. By integrating automation with human oversight, the framework addresses two critical challenges: reducing repetitive errors and ensuring compliance with quality standards such as ISO 18587. The inclusion of a pilot protocol further strengthens the practical dimension of this contribution, offering a replicable method for evaluating accuracy, fluency, and efficiency in English-Romanian medical translation.

Beyond its immediate application, the proposed workflow underscores the importance of aligning technological innovation with professional ethics and domain-specific requirements. While automation can enhance productivity, it cannot replace the expertise of qualified translators and reviewers, particularly in high-risk contexts where patient safety is paramount. Future research should focus on expanding training datasets for Romanian medical texts, refining hybrid evaluation models, and exploring the integration of predictive quality estimation tools into professional environments. These directions will help consolidate a sustainable model for medical translation that balances efficiency, reliability, and ethical responsibility.

## References

1. American Translators Association (ATA). n.d. “Certification Program: Grading Standards.” Available at https://www.atanet.org [accessed November 2025].
2. Banerjee, S., & Lavie, A. 2005. “METEOR: An automatic metric for MT evaluation with improved correlation with human judgments.” In *Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for MT and/or Summarization*. Association for Computational Linguistics. pp. 65–72.
3. Svoboda, T., Biel, Ł., & Łoboda, K. (eds.). 2017. *Quality aspects in institutional translation* (Translation and Multilingual Natural Language Processing, Vol. 8). Berlin: Language Science Press. DOI: 10.5281/zenodo.1048173. Available at: https://langsci-press.org/catalog/book/181 [accessed June 2025].

4. Crăineanu, F., & Dejica, D. 2025. “English for medical purposes: Evolution, pedagogy, and future directions” in *Scientific Bulletin of the Politehnica University of Timișoara, Transactions on Modern Languages*, 24(1), 353–366. Available at: https://dspace.upt.ro/xmlui/handle/123456789/7656.
5. Dejica, D. 2012. ‘Developing the Electronic Tools for Translators Syllabus at Politehnica University of Timisoara‘. In Procedia-Social and Behavioral Science Journal, Volume 46, 2012, Pages 3614-3618, ISSN 1877-0428. Published by ELSEVIER.
6. Dejica, D. 2016a. “Understanding technical and scientific translation: A genre-based approach.” in *Scientific Bulletin of the Politehnica University of Timișoara, Transactions on Modern Languages*, 19(1–2), 56–66. Available at: https://sc.upt.ro/attachments/article/499/02_02_Pascalau_Dejica.pdf [accessed July 2025].
7. Dejica, D. 2016b. ‘A Standards-Based Contrastive Analysis of Online and Printed Technical Translations in Romanian’. In Dejica, Daniel & Gyde Hansen, Peter Sandrini, Iulia Para (eds.) *Language in the Digital Era. Challenges and Perspectives.* De Gruyter Open: Warsaw/Berlin. pp. 14-28.
8. Dejica, D. Eugeni, C. & A. Dejica-Cartis (eds.) 2020. *Translation Studies and Information Technology - New Pathways for Researchers, Teachers and Professionals.* Timișoara: Editura Politehnica, Translation Studies Series.
9. Dejica, D., Pungă, L., Badea, G., & Vîlceanu, T. 2022. “*Cercetarea traductologică doctorală din România*”. In M. Constantinescu, D. Dejica, & T. Vîlceanu (eds.), *O istorie a traducerilor în limba română (secolul al XX-lea)*, Vol. II. București: Editura Academiei Române.
10. Directorate-General for Translation (DG Translation). 2016. *Translation tools and workflow.* Luxembourg: Publications Office of the European Union. ISBN 978-92-79-62206-9. doi:10.2782/872834.
11. do Carmo, F., Shterionov, D., Moorkens, J. et al. 2021. “A review of the state-of-the-art in automatic post-editing.” *Machine Translation* 35, 101–143. https://doi.org/10.1007/s10590-020-09252-y.
12. European Association for Machine Translation (EAMT). 1998. *Proceedings of the EAMT Workshop: Translation technology—Integration in the workflow environment.* Geneva: EAMT. Available at: https://aclanthology.org/1998.eamt-1.
13. Grigoraș, P., & Dejica, D. 2025a. “Bridging knowledge and practice: Pedagogical strategies and research directions for technical translation in the digital age”. *Studii de Știință și Cultură*, 21(3), 147–158. Available at: https://www.ceeol.com/search/article-detail?id=1370503.
14. Grigoraș, P., & Dejica, D. 2025b. “Updating the Technical Translation Curriculum: Insights, Challenges, and Practical Recommendations” in G. Garzone (ed.) *Training Translators and Interpreters Today: Perspectives and Evolutions*. Cambridge: Cambridge Scholars Publishing. (in print.)
15. House, J. 2015. *Translation Quality Assessment: Past and Present.* London: Routledge.
16. ISO 18587:2017. 2017. *Translation Services — Post-Editing of Machine Translation Output — Requirements.* Geneva: International Organization for Standardization.
17. Karakus, I. S., Strechen, I., Gupta, A., Nalaie, K., Chen, C. L., Hassett, L. C., & Barwise, A. K. 2025. “Bridging language gaps in healthcare: A systematic review of the practical implementation of neural machine translation technologies in clinical settings.” *Journal of the American Medical Informatics Association.* doi:10.1093/jamia/ocaf150.
18. Kovacs, Z., & Dejica, D. 2025. “Medical Translation in the Age of Neural Machine Translation: Standards, Assessment Models, and Romanian Perspectives.” In *Analele Universităţii Ovidius din Constanţa. Seria Filologie*, issue 2/XXXVI/2025, Constanța: Ovidius University Press.

19. Lagarda, A.-L., Alabau, V., Casacuberta, F., Silva, R., & Díaz-de-Liano, E. 2009. “Statistical post-editing of a rule-based machine translation system.” In *Proceedings of NAACL HLT 2009: Short Papers*, pp. 217–220. Boulder, CO.
20. Lommel, A. R., Burchardt, A., & Uszkoreit, H. 2013. “Multidimensional Quality Metrics (MQM): A flexible system for assessing translation quality.” In *Proceedings of Translating and the Computer 35*, pp. 121–128. London: Aslib. Available at: https://aclanthology.org/2013.tc-1.6.pdf [accessed November 2025].
21. Mali, I., & Dejica, D. 2025a. “Localization in the Romanian context: Addressing research deficits and advancing linguistic and cultural adaptation.” in *Analele Universității Ovidius din Constanța. Seria Filologie*, 36(1), 431–454. Available at: https://litere.univ-ovidius.ro/Anale/2025_volumul_1.
22. Mali, Izabella & Daniel Dejica. 2025b. 'Localization , Technology and AI: Transforming Global Communication', in Studii de Știință și Cultură, Vol. XXI, Nr. 3, Septembrie 2025, Universitatea de Vest „Vasile Goldiş” din Arad, România, pp. 226-237.
23. Medical Language Service. 2025. “Machine Translation & Post-Editing according to ISO 18587: Efficient Solutions for Medical Translations.” Medical Language Service GmbH. Available at: https://medical-language-service.com/en/machine-translation-post-editing-according-to-iso18587-efficient-solutions-for-medical-translations/ [accessed August 2025].
24. Nitzke, J., & Hansen-Schirra, S. 2021. *A short guide to post-editing.* Berlin: Language Science Press. Available at: https://langsci-press.org/catalog/book/319 [accessed September 2025].
25. Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. 2002. “BLEU: A method for automatic evaluation of machine translation.” In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 311–318.
26. Pascalău, R., & Dejica, D. 2021. “The efficient translation flow, a key to successful business.” in *Scientific Bulletin of the Politehnica University of Timișoara, Transactions on Modern Languages*, 20(1), 50–57. Available at: https://sc.upt.ro/attachments/article/499/02_02_Pascalau_Dejica.pdf. [accessed August 2025].
27. Rojas Plata, D., & Castro Sánchez, N. A. 2024. “Neural and statistical machine translation: Confronting the state of the art.” *International Journal on Natural Language Computing*, 13(5/6), 1–25. https://aircconline.com/ijnlc/V13N6/13624ijnlc01.pdf.
28. Specia, L., Scarton, C., & Paetzold, G. H. 2018. *Quality Estimation for Machine Translation.* Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers.
29. TAUS. n.d. “Dynamic Quality Framework (DQF) and Multidimensional Quality Metrics (MQM).” Available at https://www.taus.net [accessed September 2025].
30. Waddington, C. 2003. “A Positive Approach to the Assessment of Translation Errors.” *The Translator* 9(2): 271–288.
31. Zhu, Y. 2023. “Innovative research on translation quality assessment models in the new era.” *Humanities and Social Sciences*, 11(5), 176–181. Available at: https://www.sciencepublishinggroup.com/article/10.11648/j.hss.20231105.12 [accessed August 2025].