CRITICS: Critical Science Without Borders by Translation of Scientific Knowledge

Rodrigo Agerri[♠], Itziar Aldabe[♠], Elena Cabrio[♣], Mark Cieliebak [◇], Jan Deriu [◇], Mariana Flores[♠], Jurgita Kapočiūtė-Dzikienė [Φ], Dovilė Kuizinienė [Φ], Arantza Rico[Ψ], Aritz Ruiz-González[Ψ], Aitor Soroa[♠], Mantas Vaškevičius [Φ], Serena Villata[♣]

[♠]HiTZ Center - Ixa, University of the Basque Country EHU

[♣]Université Côte d'Azur, Inria, CNRS, I3S, France

[◇]Centre for Artificial Intelligence, ZHAW School of Engineering

[Φ]Faculty of Informatics, Vytautas Magnus University, Lithuania

[Ψ]Department of Mathematics, Experimental and Social Sciences Education, University of the Basque Country EHU

rodrigo.agerri@ehu.eus

Abstract

The CRITICS project addresses science accessibility and literacy through the convergence of advanced Machine Translation (MT) based on Large Language Models (LLMs) and educational technology. By leveraging MT systems specifically optimized for scientific content, educational institutions can provide accurate, culturally relevant translations of scientific materials in students' native languages, ensuring that complex scientific concepts are comprehensible while maintaining technical accuracy. Novel research on MT for scientific documents aims to break down language barriers in accessing cutting-edge research and educational materials currently only available in high-resourced languages, thereby facilitating the democratization of scientific knowledge.

1 Introduction

CRITICS is a three-year CHIST-ERA IV Cofund 2025 project funded within the topic "Science in your own language"[1]. This call addresses the translation of scientific knowledge to bridge linguistic and cultural gaps for those who must disseminate and access scientific knowledge beyond their linguistic scope.

The project is coordinated by the HiTZ Center from the University of the Basque Country EHU[2], funded by MICIU/AEI /10.13039/501100011033 and by the European Union (PCI2025-167239-2). In addition to HiTZ, the EHU team includes researchers from the Department of Mathematics, Experimental and Social Sciences Education. The consortium is also formed by the CNRS/Université Côte d'Azur in France (grant no ANR-25-CHR4-0002-02), Vytautas Magnus University (VMU) (funded by the Research Council of Lithuania, agreement No. S-CHIST-ERA-26-1), and the ZHAW Center for Artificial Intelligence in Switzerland (grant no. 20CH-1_238349).

The accessibility to scientific content in our own languages through advanced MT naturally connects to the automated generation of science education materials, where LLMs can be applied to synthesize and adapt complex scientific concepts into level-appropriate and pedagogically sound resources. Thus, CRITICS will develop and adapt LLMs to facilitate the generation of customized science education materials (Méheut and Psillos, 2004). CRITICS will mostly follow recent science education research focused on Design-Based Research (Ruiz-González et al., 2025) and consider education materials to specify the *driving problem/questions*, the learning objectives focused on competency acquisition, the scientific practices including scientific argumentation and critical thinking, and the activities to be made by the science students. The availability of machine-translated scientific knowledge will be crucial to investigating and developing LLMs for the automatic generation of appropriate education materials in the students' native languages that relate to local students' experiences (Ruiz-González et al., 2025).

2 Objectives and Work Plan

Although CRITICS' vision applies across disciplines, the project focuses on two areas: (i) nat-

[1]https://www.chistera.eu/projects-call-2025

[2]https://hitz.eus

ural sciences (biology, chemistry, and physics) and (ii) Artificial Intelligence. The former is key to competency-based assessments, while the latter poses specific translation challenges due to the continuous introduction of new terminology that may lack established equivalents in less-resourced languages (Kleidermacher and Zou, 2026; Zhang et al., 2024). CRITICS will target a diverse spectrum of languages, namely, Basque (agglutinative language isolate), Lithuanian (East Baltic, inflectional), German (West Germanic, inflectional), and French (Romance, synthetic-fusional).

Objective 1. LLM-based Machine Translation of Scientific Documents: Focused on developing and adapting open-weight LLMs for high-quality, document-level MT of scientific texts, particularly in low- and medium-resource language pairs (e.g., English–Lithuanian/Basque/French/German).

Objective 2. Argumentation serves as a fundamental mechanism in scientific discourse, facilitating the process of reaching conclusions and facilitating science literacy and critical thinking. This objective will focus on training LLMs to recognize evidence-based argumentation and identify fallacies and scientific misconceptions.

Objective 3. Automatic Assessment and Critical Thinking: Rather than merely identifying gaps in the scientific discourse, automatic assessment will also involve the generation of Critical Questions and Detailed Feedback in competency-based assessment settings (PISA style - Programme for International Student Assessment) where cross-linguistic comparability is essential.

Objective 4. Evaluation: qualitatively evaluate the generation of critical questions and automatic assessment feedback using LLMs in a way that can be compared with human-generated judgments (Calvo Figueras and Agerri, 2025).

3 Future Work

LLM-based Translation of Scientific Documents: Terminology-aware prompting and term injection have received growing attention as key techniques for addressing an open challenge in domain-specific MT: terminological inconsistency (Sabo et al., 2024; Kim et al., 2024). This is critical in scholarly translation, where inaccurate terminology can distort meaning or undermine academic precision (Kleidermacher and Zou, 2026). It should be noted that less-resourced languages have received limited attention in LLM-based MT research (Kapočiūtė-Dzikienė et al., 2025).

The use of LLM-based evaluation for text generation tasks related to MT and Critical Questions and Feedback generation in the scientific domain remains an open research problem (Calvo Figueras and Agerri, 2025). CRITICS will provide new LLM-based evaluation methods specifically tailored to the relevant features of translating MT of scientific documents and to the science teaching-related criteria of the automatically generated Critical Questions and Assessment/Feedback.

References

Calvo Figueras, Blanca and Rodrigo Agerri. 2025. Benchmarking Critical Questions Generation: A Challenging Reasoning Task for Large Language Models. In *Findings of the EMNLP 2025*, pages 5635–5652.

Kapočiūtė-Dzikienė, Jurgita, Toms Bergmanis, and Mārcis Pinnis. 2025. Localizing AI: Evaluating Open-Weight Language Models for Languages of Baltic States. In *NoDaLiDa/Baltic-HLT*, pages 287–295.

Kim, Sejoon, Mingi Sung, Jeonghwan Lee, Hyunkuk Lim, and Jorge Gimenez Perez. 2024. Efficient terminology integration for LLM-based translation in specialized domains. In *Proceedings of the Ninth Conference on Machine Translation*, pages 636–642.

Kleidermacher, Hannah Calzi and James Zou. 2026. Science across languages: assessing LLM multilingual translation of scientific papers. In *Findings of the EACL 2026*, pages 3932–3947.

Méheut, Martine and Dimitris Psillos. 2004. Teaching–learning sequences: aims and tools for science education research. *International Journal of Science Education*, 26(5):515–535.

Ruiz-González, Aritz, Arantza Rico, and Jenaro Guisasola. 2025. Learning About Sound in Initial Teacher Training: Evaluation and Redesign of a Teaching-Learning Sequence. In *Connecting Science Education with Cultural Heritage: Selected Papers from the ESERA 2023 Conference*, pages 157–171. Springer.

Sabo, Marek, Judith Klein, and Giorgio Bernardinello. 2024. Boosting machine translation with AI-powered terminology features. In *EAMT*, pages 25–26.

Zhang, Dan, Ziniu Hu, Sining Zhoubian, Zhengxiao Du, Kaiyu Yang, Zihan Wang, Yisong Yue, Yuxiao Dong, and Jie Tang. 2024. Sciinstruct: a self-reflective instruction annotated dataset for training scientific language models. *Advances in Neural Information Processing Systems*, 37:1443–1473.