Request For Proposal

[LONG]

Use case study for a technology-based scientific translation service
OPERAS AISBL
PROPOSALS DUE BY: 7 October 2022
Company Background

OPERAS is the Research Infrastructure supporting open scholarly communication in the social sciences and humanities (SSH) in the European Research Area. Its mission is to coordinate and federate resources in Europe to efficiently address the scholarly communication needs of European researchers in the field of SSH.
Project Overview

In 2020, the French Ministry of Higher Education and Research (MESR) launched the Translations and Open Science project with the aim to explore the opportunities offered by translation technologies to foster multilingualism in scholarly communication and thus help to remove language barriers according to Open Science principles.

During the initial phase of the project (2020), a first working group, made up of experts in natural language processing and translation, published a report suggesting recommendations and avenues for experimentation with a view to establishing a scientific translation service combining relevant technologies, resources and human skills.

Once developed, the scientific translation service is intended to:

- address the needs of different users, including researchers (authors and readers), readers outside the academic community, publishers of scientific texts, dissemination platforms or open archives;
- combine specialised language technologies and human skills, in particular adapted machine translation engines and in-domain language resources to support the translation process;
- be founded on the principles of open science, hence based on open-source software as well as shareable resources, and used to produce open access translations.
Project Goals

In order to follow up on recommendations and lay the foundation of the translation service, the OPERAS Research Infrastructure was commissioned by the MESR to coordinate a series of preparatory studies in the following areas:

1. **Mapping and collection of corpora**: identifying and defining the conditions for collecting and preparing corpora of bilingual scientific texts which will serve as training dataset for specialised translation engines, source data for terminology extraction, and translation memory creation.

2. **Use cases**: drafting an overview of the current translation practices in scholarly communication and defining the use cases of a technology-based scientific translation service (associated features, expected quality, editorial and technical workflows, and involved human experts).

3. **Translation output evaluation**: evaluating the output of a set of translation engines with specialised texts.

4. **Roadmap and budget projections**: making budget projections to anticipate the costs to develop and run the service.

The four preparatory studies are planned during a one-year period as of September 2022.

The present call for tenders only covers the **(2) Use cases** study.

Two additional calls will be released in the coming months for the following studies: **(3) Translation output evaluation** and **(4) Roadmap and budget projections**.

The **(1) Mapping and collection of corpora** call is open from 1 September to 7 October 2022 (details available [here](#)).
Scope of Work

The use of technologies in professional translation has virtually become an industry standard: professional translators widely use so-called CAT tools, comprehensive and user-centred translation environments that integrate translation memories, termbases, quality check features and – more and more often – machine translation. However, the landscape of translation practices in scientific publishing is too complex to easily replicate this generalisation: translation needs and workflows can vary significantly from one discipline or publisher to another, and so the tools and human expertise required.

On the other hand, the increasing availability of performing neural machine translation systems has encouraged researchers and other members of the scientific community to adopt this technology as a writing and translation assistant. This suggests that a coordinated deployment of tailored and modular translation technologies might help to establish quality multilingualism in scholarly communication by providing involved stakeholders with a collaborative framework for an informed and structured use.

Therefore, OPERAS welcomes proposals from public and private entities to study and define the possible use cases of a technology-based scientific translation service, which will be used to translate scholarly publications and associated metadata. The use case study is expected to be based on methodologies which will involve the relevant stakeholders and thus allow the development of a user-centred service: interviews, surveys, co-design workshops, focus groups, etc.

In particular, the study will address the following points:

1. Drafting an overview of the current translation practices as well as usages of assisted writing and translation tools in scholarly communication, detailed for each of the following disciplinary domains: 1) Life Sciences, 3) Physical Sciences and Engineering, 3) Social Sciences and Humanities* and with a special focus on the English-French language pair:
   - documented workflows, use cases and scenarios;
known challenges;

- foreign-language authoring and translation within the editorial workflow;
- profiles involved in the translation process;
- tools and their usages.

* Based on the ERC panel structure available [here](#).

2. **Based on the results of the first part of the task, making suggestions for a technology-based scientific translation service:**

- Expected usage scenarios and users: how the service will be used and by whom (authors, translators, publishers, dissemination platforms, researchers, students, etc.)

- Expected features and technical requirements: specifications to make the translation service relevant, open, user-centred, collaborative, modular, compatible with other systems, and based on interoperable formats.

- Editorial workflow: integration between the translation service and the editorial workflows.

- Translation quality requirements: translation quality levels to be achieved for each type of translation on a fit-for-purpose basis, from gisting to publishable quality.
Target Deliverables and Schedule

- Three interim reports presenting the main findings for the disciplinary domains indicated above (2 pages max.). The first interim report should be delivered within the first two months, the second and the third ones within the first four months.

- Final report presenting the overall results of the use case study (15 pages max.). The final report should be delivered at the end of the six-month period allocated to the project.

- A schematised presentation of the final report (in the form of a poster or a slideshow, for example).

- Participation in five workshops to be organised by OPERAS at the end of 2022 and during the first semester of 2023.

**Final Project Due:** 30 April 2023

- Bid period: 1 September to 7 October 2022

- Result notification: 18 October 2022 EOD

- Service starting date: 1 November 2022

- Expected turnaround time: 6 months
Existing Roadblocks Or Technical Issues

- Complex landscape of practices and stakeholders to be analysed and taken into consideration for the expected projections
- Strict time frame calculated to comply with the planning of the four preparatory studies
Budget Constraints

Budget range: €60,000-€80,000
Evaluation Metrics

OPERAS will evaluate bidders and proposals based on the following criteria:

- Experience in scholarly publishing editorial workflows
- Experience in translation workflow management
- Experience in user-centred design and co-design methodologies
- Achievability of deliverables
- Adequacy of requested resources and expected results
Questions Bidders Must Answer To Be Considered

Service providers are asked to submit a service proposal describing the tasks that they will be able to perform in relation to the present call for tenders during a six-month period starting from 1 November 2022.

In particular, service providers are asked to include in their response the following information:

- Detailed description of the methodology and expertise that will be used to collect, weigh and deliver relevant use case data
- Provisional planning of the service tasks
- Detailed budget
Submission Requirements

Bidders must adhere to the following guidelines to be considered:

- Only bidders who meet all 5 metrics in the evaluation section should submit a proposal.
- Proposals must be sent in by 7 October 2022. Bidders who are interested in submitting a proposal should inform Susanna Fiorini (susanna.fiorini@operas-eu.org) no later than 30 September 2022.
- Include samples and references with your proposal.
- Proposals should not be more than 4 pages. Failure to comply with this guideline will result in an automatic rejection.
- A proposed schedule must also be included and clearly expressed.
What We’re Looking For in Potential Vendors

The call is open to public and private vendors, regardless of their country of establishment.

We are particularly interested in receiving proposals from organisations with experience or interest in academia.

We attach great value to sustainable and ethical business models.

Vendors should be able to ensure smooth communication with the steering committee throughout the entire duration of the project.
Contact Information

For questions or concerns connected to this RFP, we can be reached at:

Susanna Fiorini
susanna.fiorini@operas-eu.org