

# Report by the “Translations and open science” working group

## Executive summary

According to the *Helsinki Initiative on Multilingualism in Scholarly Communication*, multilingualism enables locally relevant research to continue, helps create impact by disseminating research results in researchers’ own languages and promotes both the full diversity of scientific work and also interaction with society. However, if scientific culture is mostly conveyed in just one language then this makes it more difficult to share knowledge beyond research organisations and universities. This is contrary to the very spirit of open science, one of the fundamental principles of which is the democratisation of access to the knowledge produced by research. The Helsinki Initiative served as a basis for the approach taken by the “Translations and Open Science” working group for its report.

The group consider that translation is clearly a possible option to respond to this requirement for openness. Their objective is to identify the possible technical means to develop the multilingual dissemination of science by capitalizing on recent advances in translation technologies. This will help promote large-scale multilingualism in scientific communication. Researchers will be able to publish in the language of their choice without being in any way penalised. A new model of universal and multilingual access to scientific information will be created. However there is one essential condition for this – “human beings must remain at the core of the process while technologies contribute to optimizing work without becoming a constraint or a source of frustration for all users, whether they are part of the actual translation process or eventual readers”.

The members of the working group highlighted three elements which are essential for the success of multilingualism:

- adapting to the scientific publishing ecosystem;
- policy actions to rethink evaluation systems, metrics and funding mechanisms;
- a cultural shift among academics, researchers and lecturers to make sure the full value of non-English publications is fully recognised.

The report suggests working towards dual objectives namely:

1) to promote the dissemination of scientific production originally in French into other languages in all continents and

2) to break down language barriers for French-speaking citizens, organisations and companies who wish to access international research results.

As an initial step, the working group stresses the importance of taking a reasoned approach to translation given the sheer mass of publications. This needs to take disciplinary habits and requirements into account (among other factors) in a differentiated approach which is necessary to promote all French scientific production internationally. To achieve this, technological experiments will be run in five scientific disciplines on three language pairs (French→English, English→French and French→Spanish) because of factors linked to the availability of

significant linguistic resources and the existence of a worldwide global audience. These experiments will work with accessible publication formats, particularly metadata, abstracts and book reviews.

The group goes on to present an inventory of translation tools in two sections – one dedicated to machine translation tools and one to computer-assisted translation (CAT) tools.

Finally, the working group recommends the following short- and medium-term actions and experiments:

1. An analysis of the nature and volume of the multilingual corpora identified and a study of new possibilities for collection;
2. Processing the collected corpora to obtain test and learning bases and shared linguistic resources to aliment translation systems;
3. Evaluation of machine translation engines using the test and learning bases;
4. The organisation of study days bringing the leaders of the winning projects of the *Scientific Translations* call for projects together with other relevant actors to make proposals based on tangible findings and results;
5. To create a demonstrator to prefigure a large-scale translation process whose results could be revised to be subsequently adapted and taken up by a publishing chain;
6. To develop a guide for researchers and research institutions on machine translation, writing in other languages and “clear writing” (adapted to machine translation);
7. To study potential avenues for collaboration on the constitution of corpora in European publishing networks.

The group works within the framework of the Committee for Open Science (CoSO) in partnership with the General Delegation on the French Language and the Languages of France (DGLFLF). Its work is also a logical extension of the call for *Scientific Translations* projects launched in 2018 by the Ministry of Higher Education, Research and Innovation (MESRI). It is monitored by a steering committee made up of representatives of the partner institutions (MESRI, CoSO, DGLFLF).