ITI & CIOL Response to UK Government's Consultation on Copyright and Artificial Intelligence regarding Translation Services

Overview of the current copyright position for translations in the UK

As we understand it, translations are protected under UK copyright law as literary works in their own right, as established in the *Copyright, Designs and Patents Act 1988 (CDPA).* Section 1(1)(a) of the *CDPA* provides protection for 'original literary works,' and Section 3(1)(a) explicitly includes translations in this category.

To qualify for copyright protection, a translation must be 'original' in the sense that it represents the translator's own intellectual creation. This does not refer to the originality of the source text, but rather to the skill, judgment, and effort needed to accomplish the translation. Translators are required to use their linguistic and cultural knowledge to make creative choices in how to render the source text in the target language.

Copyright in a translation is independent of copyright in the original work. This means that:

- Permission is needed from the original copyright holder to create and publish a translation
- The original author retains their copyright in their work
- The translator owns the copyright in their translation (unless the right is ceded to the client)

In employment situations, if a translation is created as part of a translator's employment duties, the copyright typically belongs to the employer unless there is an agreement to the contrary. For freelance translators, copyright remains with the translator unless specifically assigned to the client through the terms of the contract.

Regarding machine translation and computer-assisted translation:

- Translation memories and terminology databases (termbases) may be protected as databases under 'database rights' if there has been substantial investment in obtaining, verifying, or presenting the data in the termbase
- The current law does not specifically address the copyright status of 'raw' machine translation output
- Post-edited machine translations may qualify for copyright protection if the post-editor's contributions meet the 'originality' threshold

It is worth noting that translations often involve additional contractual obligations beyond copyright, particularly regarding confidentiality and permitted use. These contractual terms may restrict what can be done with a translation even when copyright law would otherwise permit certain uses.

The full interaction between translation copyright and AI development is currently unclear in UK law, particularly regarding the use of human translations as training data for machine translation systems. This is one of the areas where new guidance or legislation may be needed, as AI technology continues to develop. We therefore welcome this UK Government consultation on copyright and artificial intelligence.

ITI & CIOL's Views

The professional translation and interpreting sector faces unique challenges regarding copyright, intellectual property rights, and professional standards as AI technology evolves. This response from the UK's leading professional bodies for translators addresses critical areas we believe require attention.

Transparency and Data Protection

The use of professional translations as training data without consent represents a serious concern for translators and within the language services industry. Many translations are created under strict confidentiality agreements and contain sensitive client information that should be protected from unauthorised use.

For example, legal translations often contain privileged information, while medical translations frequently include patient data that is protected under multiple privacy regulations. This creates a complex web of obligations that extends beyond simple copyright considerations.

We propose that AI developers should be required to provide full disclosure about their training data sources. This should include clear documentation setting out how they source their multilingual training data, whether they use professional translations, and how they obtain parallel texts (source and translation pairs). Additionally, they should disclose any use of translation memories or terminology databases (termbases), as these represent valuable intellectual property developed by language professionals over years of practice.

Asset Protection

Professional translators invest significant time and expertise in developing specialised resources such as glossaries and translation memories. These assets, built up over years of practice, represent both intellectual property and valuable professional tools. They often contain client-specific terminology and preferred phrasing that should be protected from unauthorised use in AI training.

Technical standards need to be developed that include standardised metadata fields for translation assets. These should clearly indicate copyright status and permitted uses while providing robust mechanisms for marking translations as 'not available' for AI training.

The standards must also address how to protect translation memories and termbases, which form the backbone of many professional translation practices. Importantly, any system must preserve attribution through the AI training process to ensure professional translators receive proper credit and remuneration for their work.

Workflow Classification

The translation industry now employs multiple workflows that must be clearly distinguished in any copyright framework. Human translation with AI assistance (where professional translators use AI tools to enhance their work) differs significantly from raw machine translation or post-edited machine translation. Hybrid workflows, combining human expertise with AI capabilities, represent yet another category requiring specific consideration.

Each of these workflows demands different treatment regarding copyright protection, attribution requirements, assurance of quality standards, and professional liability. The distinctions between these categories must be clearly defined to protect both professional standards and client interests.

Mandatory Labelling

To maintain professional standards and ensure appropriate use, AI-generated translations should carry clear labels indicating their production method. These labels should specify the type of translation process used, the language pairs involved, and the level of human review applied. They must also clearly state any limitations on intended use. To not do this risks a progressive degradation in AI translation quality and ultimately risks model collapse, if machine-scraped training data repeats and reinforces incorrect translations.

A labelling system could help AI developers and users make informed decisions about the appropriate use of different translation types, while protecting professional standards. It would also provide a clear framework for liability and quality assurance.

Conclusion

In conclusion the translation and interpreting sector requires specific consideration within the AI copyright framework to protect professional standards while enabling appropriate innovation. We welcome this consultation and urge the UK Government to prioritise these language sector-specific needs in developing new regulations and guidance.