

Droits de propriété intellectuelle des données d'entraînement de l'intelligence artificielle dans la traduction automatique: Défis juridiques**Dr. Mounir KHEDAR**

Département de traduction, faculté des Lettres et des langues, université de Mohammed BOUDIAF-Msila, mounir.kheddar@univ-msila.dz.

Soumis le: 11/05/2025**révisé le:** 09/09/2025**accepté le:** 10/09/2025**Résumé**

Cet article examine les droits de propriété intellectuelle liés aux données utilisées pour entraîner les modèles d'intelligence artificielle, en particulier les mémoires de traduction. Ces mémoires sont des sources essentielles pour les systèmes de traduction automatique. L'augmentation de l'utilisation de l'IA soulève des questions juridiques et éthiques, notamment concernant les textes traduits protégés par les droits d'auteur. L'article aborde les lois sur les droits de propriété intellectuelle et les défis juridiques liés à l'utilisation de ces données. Il soulève la question de savoir si l'utilisation des données traduites viole les droits des auteurs et traducteurs. Enfin, il explore l'application de l'usage équitable et des exceptions légales pour la recherche et le développement.

Mots-clés: Données, propriété intellectuelle, intelligence artificielle, traduction automatique.

Intellectual Property Rights of Training Data for Artificial Intelligence in Machine Translation: Legal Challenges**Abstract**

This article examines the intellectual property rights associated with data used to train artificial intelligence models, with a particular focus on translation memories. These memories serve as essential resources for machine translation systems. The growing use of AI raises legal and ethical concerns, especially regarding translated texts that are protected by copyright. The article discusses intellectual property legislation and the legal challenges surrounding the use of such data. It questions whether the use of translated content infringes upon the rights of authors and translators. Finally, it explores the applicability of fair use and legal exceptions in the context of research and development.

Keywords: Data, intellectual property, Artificial Intelligence, machine translation.

Auteur correspondant: Dr. Mounir KHEDAR, mounir.kheddar@univ-msila.dz.

Introduction:**1- Évolution de la traduction et apparition des mémoires de traduction:****1-1- La traduction avant et après l'ère numérique:**

Jusqu'à une époque récente, le processus de traduction reposait presque exclusivement sur l'effort intellectuel et documentaire du traducteur. Chaque traduction constituait un travail isolé (« one-shot »): le traducteur devait mener seul une recherche approfondie dans les dictionnaires, glossaires ou corpus pour trouver les équivalents linguistiques, stylistiques et culturels, sans possibilité de réutiliser ses choix pour des projets ultérieurs. Ce procédé impliquait une forte charge cognitive (mémorisation, cohérence, fatigue), allongeait les délais et augmentait les coûts.

Un tournant majeur est apparu dans les années 1980 avec l'introduction des premiers outils de traduction assistée par ordinateur (CAT), tels que Trados (1984), puis SDL ou OmegaT. Ces logiciels ont permis de stocker et de réutiliser automatiquement les équivalents validés, tout en reposant encore sur un travail humain d'indexation et de validation terminologique. Parallèlement, les entreprises et institutions ont commencé à archiver systématiquement les paires de phrases source-traduction dans des mémoires internes, afin de garantir la cohérence et d'accroître la productivité.

Initialement confinées aux réseaux privés, ces bases de données ont progressivement évolué vers des formats standardisés (TMX, TBX) et ont permis des échanges inter-entreprises. Cette pratique de collecte et de réutilisation de segments validés a réduit les erreurs, les délais et les coûts, tout en posant les fondations des systèmes modernes d'apprentissage automatique en traduction.

1-2- Défis juridiques liés au partage des données:

Avec le développement des outils commerciaux de mémoire de traduction, il est devenu possible de partager ces données entre équipes internes et externes, ce qui a soulevé des questions éthiques et juridiques concernant la propriété des données de traduction et les limites de leur utilisation — des questions toujours débattues aujourd'hui, entre autres:

Droits de propriété: Qui possède réellement les segments stockés?

Confidentialité: Comment protéger les contenus sensibles partagés?

Licences et usage: Quelles limites pour la réutilisation par des tiers?

Rémunération: Comment valoriser le travail de recherche initial?

Ces enjeux restent débattus, notamment à l'heure où ces mémoires alimentent les systèmes d'IA.

2- Traduction automatique et intelligence artificielle: Convergence et dépendance aux données:**2-1- La Traduction Automatique Neuronale:**

Avec l'émergence de la traduction automatique sous toutes ses formes (traduction automatique statistique, traduction automatique hybride ou traduction automatique neuronale), la traduction s'effectue désormais à une vitesse fulgurante et à moindre coût, en réponse aux exigences d'un monde qui avance jour après jour à grands pas vers la diversité culturelle et l'ouverture à l'autre, où la communication efficace à travers les langues et les cultures est devenue d'une importance cruciale. Le besoin pressant d'une traduction rapide, précise et efficace dans tous les domaines (communications commerciales internationales ou coopération dans la recherche académique) est ainsi apparu.

Ce qui a sans doute contribué à cela, c'est l'émergence de l'intelligence artificielle, qui constitue le moteur principal de la traduction automatique, notamment neuronale, ce qui a notablement renforcé la qualité de la traduction et de ses résultats.

L'intelligence artificielle, sur laquelle repose la traduction automatique neuronale, est continuellement entraînée à apprendre les langues et à améliorer ses traductions. Elle repose sur la prédiction de la présence d'une séquence spécifique de mots consécutifs, ce qui l'aide à comprendre la langue plus en profondeur. À l'aide d'un ensemble d'algorithmes, elle analyse intelligemment et de manière répétée d'énormes quantités de données. La traduction

automatique neuronale capte la connaissance dans la langue source et la langue cible, en plus de leur correspondance, comme faisant partie des paramètres de l'encodeur et du décodeur qui ont été appris à partir des données.

«Neural machine translation captures the knowledge of the source and target language along with their correspondences as part of the encoder and decoder parameters learned from data. With this embedded and parameterized knowledge, a trained NMT model is able to translate a new source sentence into the target language.»⁽¹⁾.

Il est devenu donc nécessaire avec l'utilisation de la traduction automatique neuronale, de réutiliser les données de traduction humaine pour entraîner et tester les modèles de traduction. L'amélioration progressive de la qualité des résultats a élargi le champ d'application de la traduction automatique, qui ne se limite plus à ce qu'on appelle la traduction de compréhension (« gisting »), mais est également utilisée comme étape intermédiaire dans la production de traductions publiables, en tant qu'« étape intermédiaire » vers la qualité finale. Les systèmes de Machine Translation (MT) ou traduction automatique ont progressivement déplacé leur dépendance d'un simple processus de « gisting » — fournir une idée générale du sens d'un texte — vers un rôle actif dans la chaîne de production de traductions de haute qualité. Pour parvenir à ce niveau de performance, il est indispensable de réinjecter dans les modèles les données issues de traductions humaines validées.

Les algorithmes de traduction automatique neuronale fonctionnent sur le principe de l'apprentissage statistique: en ingérant d'énormes volumes de paires phrase source– phrase cible, ils « apprennent » à associer des structures et des lexiques spécifiques à chaque langue. Plus le corpus d'entraînement est riche et diversifié, meilleure est la qualité des prédictions des modèles. C'est pourquoi on puise massivement dans les mémoires de traduction et autres corpus bilingues existants pour fournir à l'IA des « exemples » multiples et variés.

Aujourd'hui, les sorties de la traduction automatique neuronale atteignent un tel niveau de qualité que la traduction automatique est souvent intégrée comme « étape intermédiaire », le traducteur reçoit un texte déjà très bien formé, qu'il se contente de post-éditer —de corriger quelques erreurs ponctuelles et ajuster le style — plutôt que de retraduire de bout en bout. Ce changement réduit considérablement le temps de production tout en conservant une qualité éditoriale élevée.

Cette amélioration a élargi l'application de la traduction automatique à des secteurs exigeants (documents juridiques, communications marketing, guides techniques), où la fiabilité et la précision sont cruciales.

Le passage des modèles statistiques aux modèles de traduction neuronale a entraîné donc une augmentation considérable du volume de données requis, ce qui est conforme à la nature des modèles d'apprentissage automatique, qui reposent sur l'extraction de motifs à partir des données via des algorithmes d'induction, sans instructions codées directement. Cette évolution s'est clairement reflétée dans l'amélioration de la fluidité et de la précision des traductions automatiques, à tel point que certaines recherches récentes ont affirmé que la traduction automatique s'approche du niveau de performance humaine. Toutefois, ce saut qualitatif a suscité à son tour de réelles inquiétudes parmi les professionnels de la traduction, notamment en ce qui concerne l'avenir de la profession dans un contexte d'automatisation croissante. La question des droits liés à l'utilisation des données de traduction pour l'entraînement des systèmes automatiques a refait surface.

3- Les enjeux juridiques liés à l'utilisation des mémoires de traduction pour l'IA:

3-1- La propriété intellectuelle des données:

Dans ce contexte de controverse, on recourt souvent à des métaphores telles que « les données sont le nouveau pétrole » ou « l'or numérique », en référence à la grande valeur économique de ces ressources. Mais ces métaphores masquent une vérité fondamentale: les données de traduction ne sont pas une ressource naturelle disponible automatiquement, mais le produit d'efforts humains organisés, généralement issus de travaux de traduction professionnels. Bien que des techniques d'exploration du web puissent être utilisées pour

extraire des textes traduits d'Internet, les données de meilleure qualité et les plus utiles restent celles issues de traductions humaines vérifiées.

Les données de traduction, qui n'étaient autrefois que des outils d'assistance dans les environnements de traduction humaine, sont aujourd'hui au cœur du développement technologique de la traduction automatique. Avec cette transformation, des défis éthiques et juridiques apparaissent, nécessitant un traitement prudent, dans un contexte de véritables inquiétudes quant à l'impact de l'automatisation sur l'avenir du métier de traducteur. Par conséquent, le débat sur la propriété des données et les droits d'utilisation n'est plus purement technique, mais s'est transformé en une question centrale dans la relation entre l'humain et la technologie.

3-2-Données parallèles et corpus bilingues dans l'apprentissage automatique:

Les systèmes de traduction automatique ont besoin, pour leur entraînement et leur apprentissage, de données qui peuvent prendre plusieurs formes, mais sont généralement bilingues, composés d'une langue source et d'une langue cible. Les phrases sont appariées dans les ensembles de sorte qu'une phrase dans un ensemble corresponde à sa traduction dans l'autre. De cette manière, le système peut apprendre à produire une phrase dans la langue cible à partir d'une phrase source, qui en est la traduction. Il est également possible d'exploiter les données bilingues avec des techniques visant à améliorer le système traditionnel ou à tenter une traduction directe sans données parallèles.

«Pour l'apprentissage de systèmes de traduction automatique, nous avons besoin de corpus de données. Ils peuvent prendre plusieurs formes. Généralement, ils sont bilingues, composés d'une langue source et d'une langue cible. Les phrases de ces deux corpus sont alignées afin qu'une phrase dans un corpus corresponde à sa traduction dans l'autre. De cette façon, le système peut apprendre à générer, à partir d'une phrase source, la phrase cible qui est sa traduction. Il est aussi possible d'exploiter des données monolingues, avec des techniques pour améliorer un système classique ou en essayant directement de traduire sans données parallèles»⁽²⁾.

Parmi les données ou bases de données les plus importantes sur lesquelles s'est appuyée la traduction automatique neuronale figurent les mémoires de traduction, qui ont transformé les traductions précédemment considérées comme des tâches distinctes en un processus continu et unifié. Cela a permis d'économiser du temps, de l'argent et des efforts, tout en améliorant la précision et l'uniformité des traductions. Cette technologie a redéfini le processus de traduction à l'ère numérique en tant que base de données stockant des « segments de texte » — sous forme d'expressions, de paragraphes ou d'unités similaires — déjà traduits auparavant et réutilisables dans des projets futurs.

«Les mémoires de traduction offrent aux traducteurs et aux entreprises un certain nombre d'avantages – on pense notamment à un accroissement de la productivité et de l'uniformité dans les textes dans certains cas, à l'élimination de tâches répétitives et ennuyeuses dans d'autres»⁽³⁾.

Il est primordial de rappeler que les technologies de traduction automatique neuronale ont provoqué une transformation radicale dans le secteur de la traduction, en s'appuyant désormais sur des modèles d'apprentissage profond capables de simuler les structures linguistiques et les contextes stylistiques avec un degré de précision inédit. Pour entraîner ces modèles, les données linguistiques bilingues – en particulier les mémoires de traduction – constituent une ressource de grande valeur, en raison des segments traduits qu'elles contiennent, dont la qualité a été vérifiée et validée dans des contextes professionnels.

«However, neural networks ideally require large amounts of high- quality training data and this is particularly true of NMT since the popularization of the system architecture known as the transformer model, which has exhibited efficiency gains (Vaswani et al. 2017). As a result, human translation data has become a valuable resource. This may be illustrated by high valuations for language data brokers, who sell language data for machine learning system training, and media stories of a 'boom in language data' (Diño 2018). Some publicly available

translation data, such as that from the Directorate- General of Translation in the EU, is used as a basis for MT training, although a greater amount and breadth of data from other sources is required for an MT system to have a competitive advantage⁽⁴⁾ ».

Et en raison de l'utilisation massive des données issues des mémoires de traduction dans les processus d'entraînement de l'intelligence artificielle, une question fondamentale s'est posée: l'usage de ces données sans considération des droits de propriété constitue-t-il une violation de ces droits? Et si cela s'avère effectivement une atteinte aux droits de propriété intellectuelle, à qui appartiennent réellement ces droits sur les données largement utilisées pour entraîner les systèmes? Et Quelle protection juridique leur est offerte?

3-3- Qui détient réellement les droits de propriété intellectuelle?

L'utilisation de ces mémoires pour entraîner des systèmes d'intelligence artificielle soulève un ensemble complexe de problématiques juridiques, éthiques et techniques, qui commencent par la question de savoir qui détient le droit de disposer de ces données, passent par le problème du consentement ou de la délégation des traducteurs, et s'achèvent par le risque de perte de la valeur professionnelle de leur travail du fait de la « liquéfaction » de la traduction en données brutes exploitées sans contrepartie équitable ni reconnaissance de l'effort humain.

Plusieurs considérations entremêlées sont au cœur de cette problématique, telles que la mesure dans laquelle les mémoires de traduction peuvent bénéficier de la protection du droit d'auteur, leur éventuelle classification comme bases de données protégées, les défis liés à la protection de la vie privée si elles contiennent des données sensibles, ainsi que l'absence de modèles contractuels clairs permettant la réutilisation de ces mémoires dans un cadre d'apprentissage automatique sans enfreindre les droits des traducteurs ou des tiers.

Nous tentons donc de déconstruire cette problématique à travers une analyse du cadre juridique encadrant l'usage des données issues des mémoires de traduction pour l'entraînement de l'intelligence artificielle, en examinant les modèles de gouvernance proposés pour un partage équitable des bénéfices, et en présentant une première ébauche de modèle éthique et professionnel qui concilie l'intérêt du développement technologique avec les droits des professionnels de la langue.

3-4- Risques juridiques: Reproduction et transformation des œuvres:

La violation des droits de propriété intellectuelle se manifeste tant au niveau des entrées que des sorties de l'intelligence artificielle. Concernant les entrées, le risque réside dans l'atteinte au droit de reproduction — l'un des droits fondamentaux de la propriété intellectuelle — puisque l'intelligence artificielle a besoin de copier intégralement les données des mémoires de traduction qui lui sont fournies lors de l'apprentissage. La reproduction implique l'utilisation de l'œuvre dans sa forme originale ou modifiée, que ce soit par sa fixation sur un support matériel ou par tout moyen permettant sa copie ou l'extraction d'une partie.

La Convention de Berne stipule à son article 9 que les auteurs des œuvres littéraires et artistiques protégées par ladite convention jouissent du droit exclusif d'autoriser la reproduction de ces œuvres, sous quelque forme et par quelque procédé que ce soit.

Quant aux sorties de l'IA, elles peuvent être considérées comme des adaptations ou transformations (adaptation) des œuvres, car la créativité algorithmique s'est développée au point de pouvoir réinterpréter les styles et les schémas des écrivains et artistes dans des formes numériques, menant ainsi à la création de nouvelles œuvres fondées sur celles des auteurs originaux.

3-5- Traduction comme œuvre dérivée selon la Convention de Berne:

La Convention de Berne de 1886 — premier instrument juridique international — a expressément reconnu la traduction comme une œuvre dérivée protégée par le droit, la considérant comme une œuvre originale tant que les droits de l'auteur de l'œuvre initiale sont respectés. Elle a précisé que l'auteur détient un droit exclusif d'autoriser la traduction de son œuvre, qu'il s'agisse d'œuvres littéraires ou artistiques ou de traductions officielles de textes à caractère législatif, administratif ou juridique. Ce principe a également été confirmé par la

Convention universelle sur le droit d'auteur de Genève en 1952, puis transposé dans les législations nationales de plusieurs pays.

Les dispositions relatives aux droits de traduction ont été révisées à plusieurs reprises au cours des premières décennies de la Convention de Berne, mais les traductions ont toujours été considérées comme des œuvres dérivées. L'article 8 de ladite convention, intitulé « Droit de traduction », stipule que les auteurs des œuvres littéraires et artistiques protégées par la présente convention jouissent du droit exclusif de traduire ou d'autoriser la traduction de leurs œuvres pendant toute la durée de protection des droits sur les œuvres originales.

"Les auteurs d'œuvres littéraires et artistiques protégés par la présente Convention jouissent, pendant toute la durée de leurs droits sur l'œuvre originale, du droit exclusif de faire ou d'autoriser la traduction de leurs œuvres⁽⁵⁾."

Ces principes posés par la Convention de Berne éclairent directement la question contemporaine des mémoires de traduction (TM). Puisque chaque traduction constitue une œuvre dérivée protégée, les segments stockés dans une mémoire appartiennent en premier lieu à celui qui détient les droits sur la traduction: traducteur, agence ou client, selon les contrats en vigueur. Ainsi, l'exploitation de ces mémoires — notamment pour l'entraînement de systèmes d'intelligence artificielle — ne peut être légitime que si elle respecte le droit exclusif reconnu par la Convention, c'est-à-dire l'autorisation préalable du titulaire des droits.

4- La mémoire de traduction comme base de données: implications juridiques:

4-1- Nature technique des mémoires de traduction:

Les mémoires de traduction sont des bases de données informatiques qui enregistrent des segments de texte déjà traduits — phrases, paragraphes ou unités similaires — et leurs équivalents dans une ou plusieurs langues. Ces segments, organisés en paires appelées « unités de traduction », permettent de réutiliser les travaux antérieurs afin d'améliorer la rapidité et la qualité des traductions ultérieures.

« Elles font partie des [...] outils de rappel automatique [...] qui reposent sur la réutilisation ou le recyclage des unités linguistiques. Comme le précise L'Homme, le « type de réutilisation effectuée par les mémoires de traduction repose sur l'établissement de correspondances entre des phrases à traduire et des phrases stockées dans un répertoire » (ibid.). Règle générale, les mémoires de traduction comportent un aligneur, un logiciel de terminologie et un dépouilleur (ibid.). Dans la majorité des cas, l'unité de segmentation des mémoires de traduction est la phrase»⁽⁶⁾.

«..... Pendant que le traducteur travaille sur la traduction d'un fichier, le programme affiche chaque segment du texte original successivement et fournit une traduction antérieure à réutiliser, si le programme trouve un segment correspondant dans sa base de données.»⁽⁷⁾.

(Notre traduction)

L'introduction de ces outils au début des années 1990 a suscité des débats, l'idée de réutiliser des traductions existantes n'étant pas immédiatement acceptée. Les premiers utilisateurs ont néanmoins observé un gain notable de productivité. À la fin de la décennie, leur usage et l'application de tarifs réduits pour les segments répétés sont devenus presque incontournables dans les collaborations avec les prestataires linguistiques.

Avec leur évolution, ces outils ont gagné en performance, intégrant de vastes mémoires, des bases terminologiques complexes et des formats d'échange favorisant le travail collaboratif. Toutefois, certains traducteurs ont exprimé des réserves, notamment sur la qualité des traductions produites à partir de mémoires qu'ils n'avaient pas contribué à élaborer.

Cette situation a suscité des interrogations sur la propriété des mémoires de traduction et sur la question de savoir si elles appartiennent au traducteur, au fournisseur de services ou au client final. Dans ce contexte, des questions éthiques et juridiques ont été soulevées concernant la légitimité du partage et de l'utilisation des mémoires en dehors de leur contexte d'origine, soulignant la nécessité d'un cadre contractuel clair pour l'échange de ces données.

Du point de vue juridique, GOW⁽⁸⁾ a abordé les questions relatives aux droits d'auteur dans le contexte des bases de données, estimant que les mémoires de traduction ne remplissent pas le niveau d'« originalité » nécessaire pour être pleinement protégées en raison de leur nature fragmentée et composite composée de plusieurs segments. Bien qu'il n'y ait pas de clarté juridique, il convient de noter que la remise des fichiers de mémoire au client est devenue une pratique courante et que son refus peut entraîner des tensions dans la relation professionnelle ou empêcher le règlement des paiements dus.

Dans le cadre plus large de la propriété intellectuelle, l'Accord sur les aspects commerciaux des droits de propriété intellectuelle (TRIPS), en vigueur depuis 1995, stipule que les logiciels et bases de données sont classés parmi les œuvres littéraires selon les critères de la Convention de Berne, à condition qu'ils comportent une création intellectuelle indépendante du contenu des données. Cependant, **GOW** estime que les mémoires de traduction ne remplissent pas cette condition aux États-Unis et au Canada, ce qui affaiblit la possibilité de les protéger dans ce cadre juridique.

4-2- Législation européenne et protection sui generis:

En revanche, la Directive de l'Union européenne sur les bases de données (1996) offre une protection plus large en attribuant ce que l'on appelle des droits spécifiques (*Sui Generis*), qui peuvent être appliqués à des bases de données individuelles ou collectives. L'article 7 de la directive européenne précise que cette protection est accordée lorsqu'un investissement substantiel a été réalisé dans la collecte, l'organisation ou la présentation du contenu de la base de données, et elle dure 15 ans à partir de la dernière modification. Il est important que cet investissement soit distinct des processus de traduction eux-mêmes. La protection accorde au titulaire du droit un monopole sur la réutilisation ou l'extraction complète ou partielle de la base de données, avec des exceptions pour les fins pédagogiques, de recherche et de service.

Malgré ces cadres juridiques, la détermination de la propriété des mémoires de traduction reste une question complexe en raison de l'interconnexion des droits relatifs au texte original, à la traduction et à la structure de la mémoire elle-même.

Avec le changement radical dans le développement des systèmes de traduction automatique, avec l'émergence des modèles basés sur les données (*Data-driven Machine Translation*), qui ont surpassé les systèmes basés sur les règles linguistiques traditionnelles, ces modèles ont imposé de nouvelles exigences, dont la disponibilité de textes bilingues alignés utilisés pour entraîner les systèmes. Cela a relancé des interrogations fondamentales sur les droits des traducteurs et des utilisateurs finaux dans un contexte de développement technologique accéléré.

4-3- Enjeux de la suppression des métadonnées:

L'une des problématiques majeures réside dans l'élimination des métadonnées associées aux textes avant leur entrée dans les systèmes de traduction automatique, ces données pouvant révéler l'identité du traducteur et la date de la traduction. Bien que certains chercheurs recommandent l'utilisation de ces données pour définir l'étendue d'utilisation des traductions à des fins de formation, cette suggestion n'a pas encore été mise en œuvre comme norme. Au contraire, la suppression des métadonnées est devenue une pratique courante, partiellement motivée par les exigences légales liées à la protection des données personnelles, comme le règlement général sur la protection des données (RGPD) adopté par l'Union européenne en 2018. Ainsi, ce processus de dissimulation contribue à protéger l'identité des traducteurs, mais complique en même temps la possibilité de prouver la propriété ou de revendiquer des droits d'utilisation.

D'un point de vue technique, les outils de mémoire de traduction (*Translation Memory - TM*) reposent sur la division des textes en segments de traduction, comprenant des phrases, des titres ou des éléments de listes. Ces segments sont ensuite utilisés pour former des systèmes de traduction statistique qui s'appuient sur des tableaux de phrases composés de petites unités linguistiques. Cependant, l'augmentation de l'utilisation de la traduction automatique neuronale (*Neural Machine Translation - NMT*) depuis 2016 a marqué un

tournant majeur. Il est désormais possible de réutiliser les traductions passées à plusieurs niveaux, y compris les mots, les sous-mots (subwords) et même les caractères. Ce développement a permis de traiter des vocabulaires complexes et des termes techniques de manière plus efficace, en particulier avec l'utilisation du modèle Transformer, qui a montré des performances exceptionnelles.

5- Défis actuels de la reconnaissance juridique des contributions humaines:

5-1- Difficultés à identifier les contributions individuelles:

En raison du recours à des quantités massives de données de traduction humaine pour entraîner les modèles de traduction automatique neuronale, il est désormais pratiquement impossible de suivre la contribution d'un traducteur individuel ou de déterminer son influence créative. Ce problème se complique dans les contextes de travail collaboratif ou de traduction participative (crowdsourcing), qui ne sont soumis à aucun cadre juridique clair régissant la propriété des données ou leur utilisation.

Dans ce contexte, une question essentielle se pose: est-il possible de soumettre les résultats de la traduction automatique aux droits d'auteur? Selon la législation actuelle, comme la loi britannique de 1988, le droit d'auteur est attribué à la personne qui organise les arrangements techniques pour la production de l'œuvre informatique. Aux États-Unis, la Commission CONTU (National Commission on New Technological Uses of Copyrighted Works) de 1978 a précisé que l'auteur est celui qui exploite l'ordinateur, et non le système lui-même. La plupart des systèmes juridiques mondiaux considèrent que la créativité est attribuée uniquement à l'homme, ce qui rend la protection des résultats automatiques controversée. En conséquence, les traductions brutes produites par la machine (raw MT output) sont souvent considérées comme inéligibles à la protection juridique en raison de l'absence d'un élément créatif humain.

6-Vers un cadre juridique adapté à l'IA dans la traduction:

6-1- Limites des législations existantes:

Dans ce vide juridique, la relation entre le traducteur humain, le fournisseur de services et le développeur du système repose souvent sur des contrats individuels qui échappent à des normes unifiées, ce qui crée une ambiguïté juridique qui entrave l'élaboration d'un cadre clair pour la distribution des droits et des responsabilités. Ainsi, certaines entités, comme le Parlement européen dans sa résolution sur les règles civiles relatives aux robots (2017), recommandent l'élaboration d'un cadre juridique prenant en compte les nouvelles complexités issues de l'intelligence artificielle, non seulement en matière de droits d'auteur, mais aussi en termes de responsabilité juridique en cas de préjudice causé par une utilisation incorrecte des résultats automatiques, comme dans les traductions médicales ou juridiques.

6-2- Propriété du contenu vs propriété de la forme :

La propriété intellectuelle des mémoires de traduction, même avant leur utilisation à grande échelle comme données pour entraîner l'intelligence artificielle, est un sujet complexe, car elle implique souvent des droits croisés qui peuvent être divisés en deux parties principales: un lien au contenu ou au fond de la mémoire de traduction et l'autre à sa forme.

En ce qui concerne le contenu, deux parties principales peuvent revendiquer sa propriété:

1- Le traducteur, qu'il soit une personne physique, employée ou indépendante, ou une personne morale comme une agence de traduction ou une entreprise.

2- Le client (personne morale ou entreprise), qui fournit le texte source ou en est l'auteur.

Quant à la forme:

3- L'intermédiaire ou le créateur de la base de données.

«The number of potential owners of a translation memory can at least be restricted to three: the translator, the translation buyer (client), and any agency involved. Ownership could also be shared»⁽⁹⁾.

Mais avant de s'engager et d'approfondir ce problème juridique, il nous faut d'abord expliquer la nature de la propriété intellectuelle. La première définition donnée des droits de propriété intellectuelle est venue dans la convention de Stockholm pour la création de

l'Organisation mondiale de la propriété intellectuelle, et il est dit dans son article 2, paragraphe 8, que les droits de propriété intellectuelle.

- aux œuvres littéraires, artistiques et scientifiques, aux interprétations des artistes interprètes et aux exécutions des artistes exécutants, aux phonogrammes et aux émissions de radiodiffusion, - aux inventions dans tous les domaines de l'activité humaine, aux découvertes scientifiques, - aux dessins et modèles industriels, - aux marques de fabrique, de commerce et de service, ainsi qu'aux noms commerciaux et dénominations commerciales, à la protection contre la concurrence déloyale; et tous les autres droits afférents à l'activité intellectuelle dans les domaines industriel, scientifique, littéraire et artistique.

"Les droits de propriété intellectuelle se divisent en deux catégories: l'une concerne les droits de propriété industrielle et l'autre concerne les droits de propriété littéraire et artistique, qui est la catégorie qui nous intéresse dans cette recherche.

Et comme les mémoires de traduction sont des bases de données qui stockent des textes originaux et leurs traductions, certains de ces textes, ou tous, peuvent appartenir à l'une des œuvres mentionnées ci-dessus protégées par des droits de propriété intellectuelle et littéraire. Dans ce cas, deux parties principales ont le droit de revendiquer ces droits liés au contenu de la mémoire de traduction.

6-2-1- La propriété intellectuelle en ce qui concerne le contenu de la mémoire de traduction:

6-2-1-1- Les droits du traducteur:

Les traducteurs croient que la mémoire de traduction leur appartient exclusivement, à eux seuls, car ce sont eux qui ont effectué les recherches documentaires, les opérations de transfert, de traduction, et d'autres travaux de révision et de correction linguistique. Ils considèrent que leur expérience avec la langue, leurs compétences, et le temps qu'ils ont investi justifient leur propriété de celle-ci et, par conséquent, la considèrent comme le résultat de leur créativité intellectuelle et de leurs efforts personnels.

La Convention de Berne a explicitement reconnu les droits d'auteur des traductions, les considérant comme des œuvres dérivées qui doivent être protégées sans porter atteinte aux droits de l'œuvre originale.

Lorsque les traductions présentes dans les mémoires de traduction sont protégées par des droits de propriété intellectuelle, en tant qu'œuvres dérivées des textes originaux, les traducteurs peuvent revendiquer les droits d'auteur sur ces œuvres, et leur utilisation sans autorisation peut constituer une violation grave de ces droits.

Cependant, il convient de souligner un point important à ce sujet, à savoir que les lois sur la protection des droits de propriété intellectuelle stipulent que, pour qu'une œuvre créative puisse obtenir ces droits, elle doit être originale et d'une certaine taille. Par exemple, les mots et expressions courants ne peuvent pas être protégés, car, par définition, ils ne sont pas originaux. De même, aucun droit de propriété intellectuelle ne sera attribué à une "œuvre" qui se compose d'une seule phrase (sauf s'il s'agit d'un texte publicitaire ou d'un poème). D'une part, et d'autre part, la Convention de Berne a reconnu dans son article 10 la liberté d'inclure des citations à des fins d'illustration ou d'explication. Cela signifie qu'il est permis de faire des citations d'œuvres et documents protégés par des droits de propriété intellectuelle sans avoir à obtenir l'autorisation de l'auteur original, à condition que l'objectif de ces citations soit légitime.

«Libre utilisation des œuvres dans certains cas: 1. Citations; 2. Illustration de l'enseignement; 3. Mention de la source et de l'auteur. Autres possibilités de libre utilisation des œuvres: 1. De certains articles et de certaines œuvres radiodiffusées; 2. D'œuvres vues ou entendues au cours d'événements d'actualité»⁽¹⁰⁾.

Dans certains systèmes juridiques, le *fair use* permet la réutilisation de segments issus des mémoires de traduction sans autorisation, notamment à des fins pédagogiques ou fonctionnelles. Cette exception, bien qu'elle assouplisse la protection des œuvres, suscite des

tensions avec le droit d'auteur, garant de l'originalité et de l'exploitation économique des créations. La problématique s'étend à l'entraînement des modèles d'IA, où l'utilisation de contenus protégés peut constituer une contrefaçon difficile à prouver. Pour y répondre, certains pays ont instauré des droits spécifiques d'exploration de données. L'affaire Getty Images contre Stability AI illustre ces enjeux, Getty accusant Stable Diffusion d'avoir exploité illégalement des millions d'images protégées, entraînant une perte de revenus de licences.

"Sur cette base, étant donné que le programme de la mémoire de traduction décompose le texte original en segments textuels qui, au maximum, correspondent à une seule phrase, son contenu ne peut pas être soumis à une protection de la propriété intellectuelle en raison de sa taille non significative.

Le traducteur s'appuie sur ce point pour soutenir son point de vue selon lequel, d'un point de vue juridique, la mémoire de traduction lui appartient exclusivement, car c'est lui qui en est l'auteur, et il se base sur l'argument traditionnel connu sous le nom de 'l'effort personnel.'

6-2-1-2-Les droits de l'auteur original du texte ou du client:

Malgré ces arguments solides en faveur des traducteurs, certains experts estiment que le client ou l'auteur du texte original a également droit à la propriété de la mémoire de traduction, car bien qu'il soit vrai que le texte qu'il a fourni pour la traduction ait été décomposé en petits segments qui ne peuvent pas être soumis à une protection de la propriété intellectuelle, le texte original reste intact et complet dans le fichier de la mémoire de traduction concernée. Il est tout à fait possible de le récupérer dans sa forme initiale, et, de ce point de vue, la moitié de la mémoire de traduction appartient aux clients qui ont fourni les textes originaux".

«Sont protégés comme des œuvres originales, sans préjudice des droits de l'auteur de l'œuvre originale, les traductions, adaptations, arrangements de musique et autres transformations d'une œuvre littéraire ou artistique»⁽¹¹⁾.

En tant qu'acheteur final et donneur d'ordres pour la traduction, le client peut également revendiquer des droits sur la mémoire de traduction, car c'est lui qui demande la traduction et paie pour celle-ci. Par conséquent, il a le droit de contrôler tous les éléments issus du processus de traduction, y compris la mémoire de traduction.

Le client ou l'auteur du texte original, c'est-à-dire le document qui a été décomposé en petites unités dans la mémoire de traduction, s'il est protégé par des droits de propriété intellectuelle, peut revendiquer des droits sur la mémoire de traduction. À cet égard, la Convention de Berne a reconnu les droits moraux et financiers des auteurs originaux, et les droits moraux comprennent le droit à l'attribution et à l'intégrité (le droit de l'auteur de revendiquer la paternité de son œuvre et de s'opposer à toute modification de celle-ci).

Les droits financiers, quant à eux, incluent les droits de reproduction et de distribution des œuvres, ainsi que le droit de concéder des licences pour les œuvres dérivées, telles que les modifications, adaptations et traductions.

6-3- La propriété intellectuelle en ce qui concerne la forme de la mémoire de traduction :

Enfin, si une base de données de mémoire de traduction est créée et maintenue par une agence de traduction ayant contracté avec un traducteur ou des traducteurs externes, elle peut se considérer comme propriétaire de la mémoire de traduction, car elle a financé sa création et son entretien, et la considère comme un de ses actifs commerciaux qui renforce sa compétitivité sur le marché.

6-3-1- La protection juridique des bases de données numériques :

L'émergence des bases de données numériques a soulevé plusieurs problèmes, principalement en raison de l'insuffisance des dispositions légales traditionnelles pour protéger ce type d'œuvres à caractère technique. Cela m'amène à aborder la protection juridique des bases de données numériques en vertu des conventions internationales, puis nous aborderons la protection prévue par la directive européenne sur la protection des bases de données, en raison de son importance. En outre, nous évoquerons la protection légale

instaurée par le législateur algérien dans son système juridique, qui comprend la protection civile, l'une des principales pierres angulaires de la protection des bases de données numériques, en raison des compensations qu'elle offre à l'auteur pour réparer le préjudice subi, ainsi que la protection pénale, qui est essentielle à toute protection juridique au sein de tout système législatif.

6-3-2- Efforts internationaux pour la protection des bases de données numériques :

Face à l'importance croissante des droits de propriété intellectuelle en général et des bases de données numériques en particulier, et compte tenu de la facilité d'accès à ces dernières, ce qui les rend vulnérables aux atteintes, les pays ont cherché des solutions à ce phénomène en mettant en place un système conventionnel visant à protéger ce type de droits intellectuels, à travers l'adoption de mesures préventives et dissuasives.

Les bases de données (Databases) sont des systèmes organisés de données ou d'informations, préparées et classées selon une méthodologie spécifique, démontrant un effort intellectuel ou matériel tangible dans leur collecte, leur classification ou leur organisation, en utilisant l'ordinateur comme moyen technique pour le stockage, le traitement et la récupération.

Les bases de données qui témoignent de créativité intellectuelle méritent d'être protégées par les droits d'auteur, et la protection couvre à la fois les données elles-mêmes et la structure organisationnelle (le contenant) qui rend la base de données une œuvre unique. La protection juridique des bases de données ne se limite pas au simple stockage des données, mais est directement liée à l'élément d'innovation ou d'effort déployé pour leur organisation, tant en termes de contenu, de structure technique que de moyens d'accès et de récupération.

C'est pourquoi plusieurs conventions internationales ont été mises en place pour tenter d'inclure un nombre plus important de pays et rendre le système juridique de protection des droits d'auteur un système fondé sur la coopération internationale visant à offrir la meilleure protection possible à ces droits. Cela a commencé par la Convention de Berne pour les œuvres littéraires et artistiques, signée le 19 septembre 1886, qui a été complétée le 4 mai 1896 à Paris, modifiée lors de la conférence de Berlin le 13 novembre 1908, à la conférence de Rome le 2 juin 1928, à Bruxelles le 26 juin 1948, puis à Stockholm le 14 juillet 1967, à Paris le 24 juillet 1971, et modifiée le 28 septembre 1979."

En plus de la Convention de Berne, plusieurs autres conventions ont émergé, telles que la Convention de Genève, connue sous le nom de Convention universelle sur le droit d'auteur, signée le 6 septembre 1952, ainsi que la convention établissant l'Organisation mondiale de la propriété intellectuelle à Stockholm le 14 juillet 1967.

Bien que la Convention de Berne soit la plus ancienne, elle est devenue incapable de fournir une protection juridique pour ce nouveau type d'œuvre numérique. En réponse à ce besoin, la Convention sur les aspects commerciaux des droits de propriété intellectuelle 'TRIPS' a été émise, et elle est considérée comme la première convention à aborder directement la protection des bases de données sous leur forme numérique. Ensuite, il y a eu le Traité de l'OMPI (Organisation Mondiale de la Propriété Intellectuelle) sur la protection des droits d'auteur de 1996, également connu sous le nom de 'Premier traité Internet', qui a également abordé la protection des bases de données à travers plusieurs dispositions qui montrent son intérêt pour le développement technologique et ses implications pour les droits des auteurs et leurs œuvres dans l'environnement numérique.

Le Traité de l'OMPI de 1966 sur les droits d'auteur précise que les bases de données sont des créations intellectuelles qui doivent être protégées en tant que telles."

Compilations de données (bases de données)

Les compilations de données ou d'autres éléments, sous quelque forme que ce soit, qui, par le choix ou la disposition des matières, constituent des créations intellectuelles sont protégées comme telles. Cette protection ne s'étend pas aux données ou éléments eux-mêmes et elle est sans préjudice de tout droit d'auteur existant sur les données ou éléments contenus dans la compilation⁽¹²⁾.

Les conventions internationales ont confirmé cette tendance, car l'article 10(2) de l'Accord TRIPS stipule que Les compilations de données ou d'autres éléments, qu'elles soient reproduites sur support exploitable par machine ou sous toute autre forme, qui, par le choix ou la disposition des matières, constituent des créations intellectuelles seront protégées comme telles. Cette protection, qui ne s'étendra pas aux données ou éléments eux-mêmes, sera sans préjudice de tout droit d'auteur subsistant pour les données ou éléments eux-mêmes.

Bien qu'il existe des efforts internationaux sérieux pour protéger les bases de données, ils restent insuffisants pour protéger les droits matériels des investisseurs dans les bases de données qui ont effectué un investissement substantiel, face aux droits accordés à l'auteur en vertu de la législation sur le droit d'auteur, qu'il s'agisse de droits matériels ou moraux, ainsi que des exceptions en découlant, qui ont constitué des obstacles pour les producteurs et investisseurs de bases de données. Cela a conduit à l'adoption de la directive européenne sur la protection des bases de données en 1996.

6-3-2-1- La protection juridique des bases de données numériques selon la directive européenne :

Malgré les efforts internationaux déployés pour offrir une protection juridique efficace aux auteurs et à leurs œuvres, qu'elles soient classiques ou numériques, la plupart de ces accords restent traditionnels et ne suivent pas le rythme rapide du développement technologique, notamment en ce qui concerne la protection des œuvres numériques qui sont désormais exposées à diverses attaques électroniques, étant donné l'importance croissante des bases de données dans divers domaines, notamment scientifiques et économiques.

Etant donné les lacunes dans la protection juridique des bases de données dans diverses législations comparées, notamment parce que les législations nationales connaissent une certaine instabilité quant au cadre juridique des bases de données, cela a poussé le Parlement européen et le Conseil de l'Union européenne à chercher un système juridique unifié. Ces efforts ont abouti à l'adoption de la directive du Parlement européen et du Conseil européen n° 9/96/CE en date du 1er mars 1996 concernant la protection juridique des bases de données."

"À travers l'étude de la directive européenne concernant la protection des bases de données, il apparaît que le contenu de la protection des œuvres de bases de données diffère entre la protection par droit d'auteur et la protection par ce que cette directive appelle le 'droit sui generis'.

Cependant, la nature internationale d'Internet a conduit à la facilité d'accomplir des crimes sur les bases de données numériques dans le cas où les deux parties impliquées appartiennent à un pays donné, tandis que le résultat de cet acte criminel se réalise dans un autre pays, ce qui soulève des difficultés pour déterminer la législation applicable entre les lois des différents pays.

Par conséquent, les autorités d'enquête dans un pays donné ne peuvent plus faire face à ce type de crime seul, sans l'aide ou la collaboration d'autres autorités dans d'autres pays, ce qui appelle la nécessité d'une coopération dans la lutte contre les crimes liés à Internet, notamment les atteintes aux bases de données.

7- La protection juridique des mémoires de traduction :

Les mémoires de traduction sont l'une des applications avancées des systèmes de calcul linguistique. Elles se composent d'unités bilingues (segments) dans lesquelles sont stockées des phrases ou des expressions déjà traduites, de sorte qu'elles sont récupérées automatiquement lorsqu'une phrase similaire ou identique apparaît pendant un processus de traduction futur. Par conséquent, elles adoptent une structure de base claire, comprenant un regroupement organisé des unités de données linguistiques qui peuvent être traitées et récupérées automatiquement, ce qui les place du point de vue conceptuel dans la catégorie des bases de données.

Il apparaît clairement qu'une mémoire de traduction, lorsqu'on la considère comme une base de données, est protégée par la législation sur les droits des bases de données. En effet, la structure et l'organisation des données dans la mémoire peuvent être protégées, ce qui

pourrait nécessiter une autorisation préalable pour toute extraction ou réutilisation importante de ces données.

Cependant, en principe, les données linguistiques brutes dans la mémoire de traduction – comme une phrase traduite ou un mot unique – ne sont pas protégées en soi, à moins qu'elles n'impliquent un certain degré de créativité ou d'effort personnel, ce qui est conforme au principe établi dans l'Accord TRIPS et la jurisprudence internationale. Cependant, l'effort consenti pour rassembler ces unités et les organiser dans une structure récupérable automatiquement, avec l'utilisation d'outils de classification linguistique (tels que les temps, les sujets, les paires linguistiques, etc.), peut atteindre le niveau de "l'innovation structurelle" ou de "l'investissement substantiel", permettant ainsi de soumettre ces mémoires au système de protection des bases de données.

Ainsi, on peut considérer les mémoires de traduction comme des bases de données. Dans ce cas, pour qu'une base de données bénéficie de la protection de la propriété intellectuelle, elle doit répondre à certaines conditions, dont la condition d'originalité qui découle d'un effort créatif inhérent à la méthode d'expression et à l'unicité de l'auteur. L'originalité de la base de données réside dans sa forme et sa structure, car son contenu ne représente souvent aucune originalité étant donné qu'il s'agit d'œuvres d'autrui. En l'absence de l'originalité requise, la base de données peut apparaître soit comme un mélange d'œuvres protégées préexistantes qu'aucun ne peut revendiquer comme siennes, soit comme un amalgame de données brutes qui ne sont la propriété de personne et sont accessibles à tous, comme les chiffres."

Conclusion:

La question de la propriété et de l'exploitation des mémoires de traduction demeure entourée d'incertitudes juridiques, notamment en matière de droits de propriété intellectuelle. En pratique, plusieurs scénarios de responsabilité coexistent.

L'auteur du texte source reste titulaire des droits initiaux, dès lors que son œuvre présente un caractère original.

Le traducteur, pour sa part, peut revendiquer des droits d'auteur sur la traduction considérée comme une œuvre dérivée, à condition qu'elle témoigne d'un apport créatif.

Le client ou commanditaire de la traduction, en vertu du contrat conclu, peut disposer de droits patrimoniaux ou d'un usage exclusif des traductions produites.

Le concepteur ou gestionnaire de la base de données de traduction bénéficie, lui, de droits spécifiques liés à la structuration, à l'indexation et à l'alignement des unités de traduction.

Enfin, dans le cas de bases alimentées par des projets antérieurs, la propriété se complexifie, car chaque auteur, chaque traducteur et parfois même les prestataires techniques (programmeurs, fournisseurs d'outils) peuvent prétendre à des droits ou à une reconnaissance de leur contribution.

«The situation is further complicated by the inclusion of legacy data from previous projects. This introduces ownership claims involving each author of the source content and each translator of the target content already in the database, as well as potentially the language service provider or client through database rights covering the alignment of the translation units»⁽¹³⁾.

Cette pluralité d'acteurs génère un chevauchement des revendications qui complique l'attribution des droits et la répartition des responsabilités, surtout lorsque ces mémoires sont utilisées pour entraîner des systèmes d'intelligence artificielle.

Face à ces enjeux, il existe une véritable urgence législative et réglementaire. Plusieurs axes apparaissent nécessaires: d'abord, une clarification juridique visant à définir précisément le statut des mémoires de traduction et à identifier les titulaires légitimes des droits; ensuite, une adaptation du droit d'auteur afin de prendre en compte les usages liés à l'intelligence artificielle, en particulier l'entraînement des modèles sur des corpus bilingues protégés; également, un encadrement des exceptions comme le fair use et le text and data mining, afin de favoriser la recherche et l'innovation sans porter atteinte aux intérêts des créateurs; enfin, la mise en place d'un régime de responsabilité partagé qui établisse des règles équilibrées,

garantissant à la fois la reconnaissance et la rémunération des différents acteurs et une sécurité juridique pour les développeurs et utilisateurs d'IA.

En somme, l'avenir des mémoires de traduction et leur intégration dans les systèmes d'apprentissage automatique ne peut être envisagé sans une régulation claire et adaptée, capable de concilier la valorisation des innovations technologiques avec la protection des droits des créateurs et des traducteurs.

Références:

- 1- QIAN Cao, SHAOHUI Kuang and DEYI Xiong.(2020). Learning to Reuse Translations: Guiding Neural Machine. 24th EuropeanConference on Artificial Intelligence - ECAI 2020, Santiago de Compostela, Spain.
- 2- BARDET, Adrien. (2021). Architectures neuronales multilingues pour le traitement automatique des langues naturelles, p 10.
- 3- LEBLANC, Matthieu. (2016). Les mémoires de traduction et le rapport au texte: ce qu'en disent les traducteurs professionnels. Traduction, terminologie, rédaction. TTR, 2016, Vol. 27, 02, p 128.
- 4- MOORKENS, Joss et Dave, Lewis. (2020). Copyright and the re-use of translation as data. In The Routledge Handbook of Translation and Technology. UK: Routledge, p 49.
- 5- Convention de Berne. www.wipo.int. WIPO. \[En ligne] \[Citation:04mars2024.]https://www.wipo.int/wipolex/en/text/283699.
- 6- LEBLANC, Matthieu. (2016). Op.cit. P 126
- 7- خماري سليم، أحمد. (2008). ثورة تقنية ذاكرة الترجمة. طرابلس: دار المعارف الليبية.
- 8- Voir Gow, Francie. (2007). 'You must remember this: the copyright conundrum of "translation memory"' databases'. Canadian Journal of Law and Technology, Vol. 6, 3.
- 9- Smith, Ross. (2009). Copyright Issues in Translation Memory Ownership. PricewaterhouseCoopers Spain: s.n. p 4
- 10- Convention de Berne. www.wipo.int. WIPO. \[En ligne] \[Citation: 04 mars 2024.] op.cit.
- 11- idem
- 12- OMPI. www.wipo.int. OMP. \[En ligne] \[Citation: 01 mars 2024.] https://www.wipo.int/wipolex/fr/text/295168#P69_6750[https://www.wipo.int/wipolex/fr/text/295168#P69_6750].
- 13- MAARIT Koponen, SANNA Nyqvist, and KristiinaTaivalkoski-Shilov. (2022). Translating with technology: How digitalisation affects authorship and copyright of literary texts. Using technologies for creative-text translation. NEW YORK AND LONDON: Taylor & Francis Group.p212

Bibliographie:

Livres:

- MAARIT Koponen, SannaNyqvist, and KristiinaTaivalkoski-Shilov. (2022). Translating with technology: How digitalisation affects authorship and copyright of literary texts. Using technologies for creative-text translation. NEW YORK AND LONDON: Taylor & Francis Group.
- MOORKENS, Joss et Dave, Lewis. (2020). Copyright and the re-use of translation as data. In The Routledge Handbook of Translation and Technology. UK: Routledge,
- TROUSSEL Jean-Christophe et DEBUSSCHE Julien, (2014)., Studies on translation and multilingualism: Translation and intellectual property, publication office of European union
- خماري سليم، أحمد. (2008). ثورة تقنية ذاكرة الترجمة. طرابلس: دار المعارف الليبية.

Articles:

- BARDET, Adrien. (2021). Architectures neuronales multilingues pour le traitement automatique des langues naturelles.
- GOW, Francie. (2007). 'You must remember this: the copyright conundrum of "translation memory"' databases'. Canadian Journal of Law and Technology, Vol. 6, 3.
- LEBLANC, Matthieu. 2016. Les mémoires de traduction et le rapport au texte: ce qu'en disent les traducteurs professionnels. Traduction, terminologie, rédaction. TTR, 2016, Vol. 27, 02.
- SMITH, Ross. (2009). Copyright Issues in Translation Memory Ownership. PricewaterhouseCoopers Spain: s.n.

Sites web:

- Convention de Berne. www.wipo.int. WIPO. \[En ligne] \[Citation: 04 mars 2024.] https://www.wipo.int/wipolex/en/text/283699. consulté le 04/03/2025 à 10:30.
- OMPI. www.wipo.int. OMP. \[En ligne] \[Citation: 01 mars 2024.] https://www.wipo.int/wipolex/fr/text/295168#P69_6750(https://www.wipo.int/wipolex/fr/text/295168#P69_6750).consulté le 04/03/2025. 10:45.

Conférences:

- QIAN Cao, SHAOHUI Kuang and DEYI Xiong.(2020). Learning to Reuse Translations: Guiding Neural Machine. 24th European Conference on Artificial Intelligence - ECAI 2020, Santiago de Compostela, Spain.