

Groupe de travail sur le développement juridique du système de La Haye concernant l'enregistrement international des dessins et modèles industriels

Onzième session
Genève, 12 – 14 décembre 2022

ÉTUDE SUR LA DISPONIBILITÉ ET LA FACILITÉ D'UTILISATION DES TECHNOLOGIES DE TRADUCTION

Document établi par le Bureau international

I. CONTEXTE

1. À sa dixième session tenue en 2021, le Groupe de travail sur le développement juridique du système de La Haye concernant l'enregistrement international des dessins et modèles industriels (ci-après dénommé "groupe de travail") a examiné les documents H/LD/WG/10/4 et H/LD/WG/10/5, intitulés respectivement "Version actualisée de l'étude approfondie des incidences financières et de la faisabilité technique de l'introduction du chinois et du russe dans le système de La Haye" et "Informations actualisées sur les critères de sélection des langues supplémentaires à introduire dans le système de La Haye". Le groupe de travail a demandé au Secrétariat d'établir une étude approfondie des incidences financières et de la faisabilité technique de l'introduction de nouvelles langues dans le système de La Haye, tenant compte des solutions techniques, en particulier la traduction automatique, pour examen par le groupe de travail à sa prochaine session (voir le paragraphe 18.ii) du document H/LD/WG/10/6).

2. Le présent document contient des informations détaillées sur les technologies et les pratiques de traduction utilisées par les différents services de traduction de l'OMPI. Ces informations visent à faciliter les réflexions du groupe de travail en ce qui concerne la possibilité de réduire, grâce aux technologies de traduction, les coûts liés à l'introduction de nouvelles langues dans le système de La Haye.

II. TECHNOLOGIES DE TRADUCTION UTILISÉES À L'OMPI

3. Les travaux de traduction à l'OMPI sont effectués par les différents services mondiaux de protection de la propriété intellectuelle ainsi que par la Division linguistique. Le processus de traduction se déroule généralement en deux étapes : dans un premier temps, le texte à traduire est comparé aux entrées contenues dans la mémoire de traduction applicable, puis un traducteur humain postérite les propositions de traduction générées par un modèle de traduction automatique neuronale (WIPO Translate) si un modèle est disponible dans ce domaine et pour cette combinaison linguistique, ou traduit le texte restant dans la langue cible. Enfin, une partie des traductions est soumise à un contrôle de qualité effectué par un traducteur-réviseur interne.

WORLDSERVER

4. WorldServer est un système de gestion de la traduction doté d'un outil de traduction assistée par ordinateur, qui gère le flux de travail de traduction en automatisant les tâches de traduction. Il remplit d'abord les traductions avec des segments dans la langue cible récupérés dans les mémoires de traduction, puis les segments restants avec des propositions de traduction automatique neuronale. La Section de la traduction et de la terminologie du Service d'enregistrement de Madrid et la Division de la traduction du PCT utilisent toutes deux WorldServer¹.

MÉMOIRES DE TRADUCTION

5. Les mémoires de traduction, qui constituent l'épine dorsale sur laquelle reposent les différentes pratiques de traduction, sont des recueils de textes précédemment traduits dans le domaine concerné (La Haye, Madrid, PCT ou Division linguistique). Toutes les entrées des mémoires de traduction font l'objet d'une validation par un humain pour en garantir la qualité. Les mémoires de traduction sont conservées séparément pour chaque service et chaque combinaison linguistique². Le volume des entrées contenues dans les mémoires de traduction respectives varie en fonction de la quantité de textes précédemment traduits. Par exemple, depuis 2016, la part des demandes déposées en anglais dans le cadre du système de La Haye était supérieure à 86%, celle des demandes déposées en français était d'environ 12% et celle des demandes déposées en espagnol était d'environ 1%³. Par conséquent, les mémoires de traduction pour les textes relatifs au système de La Haye dont l'anglais est la langue source comportent beaucoup plus d'entrées que les mémoires de traduction pour les textes dont le français est la langue source, tandis que les mémoires de traduction pour les textes dont l'espagnol est la langue source ne comportent pratiquement aucune entrée. Les mémoires de traduction sont intégrées dans le système de gestion des traductions (WorldServer).

WIPO TRANSLATE

6. WIPO Translate est un outil de traduction automatique fondé sur l'intelligence artificielle, mis au point en interne par l'OMPI et entraîné avec des sujets spécifiques, qui utilise la dernière technologie de traduction automatique neuronale. La traduction automatique neuronale est une technologie fondée sur des modèles de réseaux neuronaux qui "apprennent" à partir de grands volumes de textes précédemment traduits⁴. WIPO Translate a initialement été entraîné avec

¹ La Division linguistique passera à WorldServer dans un avenir proche.

² Par exemple, il existe actuellement six mémoires de traduction distinctes pour les trois langues de travail du système de La Haye (anglais-français, anglais-espagnol, français-anglais, français-espagnol, espagnol-anglais et espagnol-français).

³ Voir le graphique n° 13 de la *Revue annuelle du système de La Haye 2022*.

⁴ Par exemple, les ensembles de données utilisés pour entraîner WIPO Translate afin qu'il traduise des documents de brevet comptent généralement des millions de documents.

des documents de brevet⁵ dans les 10 langues de publication du PCT⁶ et est axé sur ces documents. Mis à la disposition des utilisateurs dans Patentscope, il fournit des propositions de traduction de ces langues vers l'anglais et inversement.

7. Un modèle sur mesure de WIPO Translate a été élaboré pour chacun des services de traduction de l'OMPI⁷. Cependant, afin de produire des traductions de haute qualité, WIPO Translate doit être entraîné avec de grands volumes de texte, dans le domaine pertinent, pour chaque combinaison linguistique⁸.

8. En outre, les propositions de traduction qui sont générées par un modèle de traduction automatique neuronale pour des combinaisons linguistiques plus éloignées sont généralement de moindre qualité que dans le cas de combinaisons linguistiques plus proches. Par conséquent, pour les combinaisons linguistiques plus éloignées, il est d'autant plus important de disposer de suffisamment de données pour entraîner le système afin de garantir une qualité de traduction qui soit satisfaisante⁹.

TRADUCTION, POSTÉDITION ET CONTRÔLE DE QUALITÉ

9. Les outils décrits ci-dessus ne permettent pas d'automatiser entièrement le processus de traduction. Ainsi, même si l'on utilise des technologies de traduction automatique, des ressources humaines restent nécessaires pour la postédition et le contrôle de qualité. Ces traducteurs doivent posséder des compétences linguistiques et en matière de traduction dans les langues respectives. À cet égard, il est rappelé qu'il est toujours préférable de confier la postédition et le contrôle de la qualité des traductions vers la langue cible à un traducteur dont la langue maternelle est la langue cible. Le contrôle de la qualité est effectué par un traducteur-réviseur interne.

10. La qualité de la traduction automatique neuronale a une incidence sur les coûts de la postédition. Si la qualité de la traduction automatique est médiocre, les coûts de la postédition de bout en bout peuvent être proches des coûts de la traduction directe par un humain, voire les dépasser.

III. PRATIQUES EN MATIÈRE DE TRADUCTION AUTOMATIQUE À L'OMPI

SYSTÈME DE LA HAYE

11. Il est rappelé que la traduction des demandes internationales dans le cadre du système de La Haye est actuellement effectuée par la Section de la traduction et de la terminologie du Service d'enregistrement de Madrid¹⁰. Les opérations de traduction relatives aux demandes dans le cadre du système de La Haye concernent l'indication du produit, la description des

⁵ WIPO Translate est entraîné avec le texte contenu dans les demandes du PCT et les collections de brevets des offices nationaux ou régionaux disponibles dans Patentscope.

⁶ Les 10 langues sont le français, l'allemand, l'anglais, l'arabe, le chinois, le coréen, l'espagnol, le japonais, le portugais et le russe.

⁷ Un modèle sur mesure est actuellement utilisé pour les traductions concernant le système de La Haye, le système de Madrid, certaines combinaisons linguistiques dans le cadre du PCT et pour certains documents de la Division linguistique.

⁸ Bien que WIPO Translate puisse potentiellement être entraîné pour traduire uniquement dans un sens, il est généralement entraîné pour traduire dans les deux sens, car en entraînant WIPO Translate avec des données parallèles (deux langues pour le même contenu) on améliore considérablement la qualité de la traduction.

⁹ Par exemple, l'anglais et le français sont des langues linguistiquement similaires, alors que l'anglais et l'arabe, le chinois ou le coréen sont des langues éloignées. L'anglais et le russe font tous deux partie de la famille des langues indo-européennes. Ils présentent toutefois des différences grammaticales, ce qui explique que la traduction d'un texte libre nécessite davantage de données pour entraîner le système.

¹⁰ Toutes les traductions (postédition) des textes relatifs au système de La Haye sont actuellement réalisées en interne.

éléments caractéristiques du dessin ou modèle industriel, les revendications et les légendes¹¹. Alors que les indications concernant le produit, les revendications et les légendes contiennent généralement moins de 10 mots, une description se compose de texte libre, avec une moyenne de 64 mots en 2022¹². À noter que le nombre de demandes internationales contenant une description est passé de 17% en 2012 à 65% en 2022¹³.

12. La Section de la traduction et de la terminologie, qui traduit directement les textes, s'appuie sur six mémoires de traduction relatives au système de La Haye pour les trois langues de travail du système de La Haye, ainsi que sur WIPO Translate. Les mémoires de traduction contenant les textes précédemment traduits relatifs aux demandes déposées dans le cadre du système de La Haye comptent actuellement un total d'environ 280 000 entrées, la majorité étant contenue dans les mémoires de traduction dont l'anglais est la langue source. Le Bureau international estime que, en moyenne, pour les traductions de l'anglais vers le français ou l'espagnol, une correspondance est trouvée dans les mémoires de traduction pour environ 35% des segments de texte, tandis que le reste du texte (65%) nécessite une postédition des propositions de traduction générées par WIPO Translate. En comparaison, pour les traductions dont la langue source est le français ou l'espagnol, le nombre de correspondances dans les mémoires de traduction est beaucoup plus faible¹⁴. Par conséquent, pour ces traductions, la majeure partie du texte fait l'objet d'une postédition à partir d'un nombre limité de propositions de traduction générées par WIPO Translate.

13. En outre, on peut observer que, bien que les mémoires de traduction ne cessent de croître, le taux de correspondance augmente beaucoup plus lentement. Le Bureau international estime que pour les textes relatifs au système de La Haye, le taux de correspondance dans les mémoires de traduction concernant le système de La Haye pourrait ne pas dépasser 40%, compte tenu de l'évolution de la situation en ce qui concerne les dessins et modèles industriels et le texte connexe contenu dans les demandes relatives au système de La Haye (en particulier les descriptions en texte libre)¹⁵. Par conséquent, le Bureau international part du principe qu'environ 60% des traductions nécessiteront encore une postédition par un humain des propositions de traduction générées par WIPO Translate.

¹¹ Voir l'article 5.2)b)ii) et iii) de l'Acte de Genève (1999), les règles 7.3)iv), 4)b) et 5)a), et 11.2) et 3) du règlement d'exécution commun à l'Acte de 1999 et l'Acte de 1960 de l'Arrangement de La Haye (ci-après dénommé "règlement d'exécution commun"), les sections 403 et 405.c) des instructions administratives pour l'application de l'Arrangement de La Haye.

¹² La longueur d'une description est très variable. Par exemple, en 2021 ainsi que dans les trois premiers trimestres de 2022, 15% des descriptions contenaient plus de 100 mots, tandis que 1% en contenait plus de 500. En 2021, 589 demandes déposées dans le cadre du système de La Haye contenaient une description d'une longueur comprise entre 100 et 500 mots, pour 30 demandes contenant une description de plus de 500 mots.

¹³ Cette évolution résulte de l'adhésion de parties contractantes dont les offices sont des offices procédant à un examen qui exigent ou recommandent d'inclure une description. À cet égard, il convient de noter que la Chine a fait une déclaration dans laquelle elle exige qu'une demande internationale contienne une brève description des éléments caractéristiques des dessins ou modèles, et que l'Office de la Fédération de Russie (ROSPATENT) recommande vivement de fournir une brève description des éléments caractéristiques du dessin ou modèle industriel.

¹⁴ En général, une correspondance est trouvée uniquement lorsqu'un même déposant dépose plusieurs demandes pour des dessins ou modèles très similaires. Dans ce cas, on trouvera davantage de correspondances pour les indications concernant le produit ou les légendes pour les descriptions.

¹⁵ Cette évolution s'explique par la nature de la protection des dessins et modèles industriels, qui exige nouveauté et originalité dans la plupart des législations nationales. L'enregistrement des dessins et modèles résulte de la mise au point de nouveaux produits donnant lieu à de nouveaux dessins et modèles, et de la création de nouveaux dessins et modèles pour des produits émergents et existants.

SYSTÈME DE MADRID

14. Il convient de rappeler que le système de Madrid et le système de La Haye ont le même régime linguistique, avec les mêmes trois langues. La majeure partie du travail de traduction effectué dans le cadre du système de Madrid concerne la traduction des indications de produits et de services dans les demandes internationales. Une liste de termes normalisés regroupe généralement lesdits produits et services; ils diffèrent donc des descriptions figurant dans les demandes selon le système de La Haye.

15. La Section de la traduction et de la terminologie du Service d'enregistrement de Madrid s'appuie sur les mémoires de traduction alimentées par les documents du système de Madrid précédemment traduits¹⁶ et sur WIPO Translate. Au total, les six mémoires de traduction de Madrid contiennent plus de six millions d'entrées. Le taux de concordance entre un document de Madrid et les mémoires de traduction de Madrid est de près de 65%¹⁷, le pourcentage restant nécessitant la postédition des propositions de traduction de WIPO Translate. La qualité des propositions de traduction générées par WIPO Translate, en particulier des traductions dont l'anglais est la langue source, permet au Bureau international de réduire les coûts de la postédition de moitié¹⁸.

16. Dans le cadre des discussions relatives à une éventuelle extension du régime linguistique, le Groupe de travail sur le développement juridique du système de Madrid concernant l'enregistrement international des marques (ci-après dénommé "groupe de travail de Madrid") envisage actuellement plusieurs mesures possibles, parmi lesquelles la création de mémoires de traduction avec un grand nombre d'équivalences terminologiques en arabe, en chinois et en russe, l'entraînement de WIPO Translate de façon à ce qu'il puisse traiter les traductions dans le cadre du système de Madrid dans ces langues et la constitution d'un corpus de listes de produits et de services traduites que les offices des parties contractantes pourraient avoir en leur possession¹⁹.

17. Toutefois, les différents types de texte contenus dans les demandes déposées selon les systèmes de La Haye et de Madrid (descriptions en texte libre par opposition à des termes normalisés), l'évolution constante du domaine des dessins et modèles industriels, l'écart entre le nombre d'entrées des mémoires de traduction des systèmes de La Haye et de Madrid (280 000 et 6 millions, respectivement) et par conséquent, la différence de taux de concordance de ces mémoires de traduction, ainsi que de la quantité de données d'apprentissages éventuelles pour chaque modèle WIPO Translate, rendent l'examen de mesures similaires moins prometteur dans le contexte du système de La Haye.

SYSTÈME DU PCT

18. La Division de la traduction du PCT est responsable de la traduction des titres et des abrégés des demandes selon le PCT des 10 langues de publication du PCT vers l'anglais et le français (s'ils n'existent pas dans ces deux langues), et de la traduction des rapports de recherche internationale et des rapports préliminaires internationaux sur la brevetabilité en anglais (s'ils n'ont pas été établis en anglais)²⁰.

¹⁶ Le système de Madrid et le système de La Haye ayant le même régime linguistique, la Section de la traduction et de la terminologie de Madrid utilise les six mémoires de traduction pour les mêmes combinaisons de langue.

¹⁷ Voir le document MM/LD/WG/19/7.

¹⁸ Cela signifie que les coûts de la postédition s'élèvent à 50% des coûts lorsque le texte a été entièrement traduit sans l'aide de WIPO Translate.

¹⁹ Voir les paragraphes 39 à 50 du document MM/LD/WG/19/7.

²⁰ Voir les règles 45.1, 48.3.c), 62*bis*.1, 72.1 et 86.2 du règlement d'exécution du Traité de coopération en matière de brevets (ci-après dénommé "règlement d'exécution du PCT").

19. Le titre d'une invention est une description brève et précise de celle-ci et il est recommandé qu'il soit compris entre deux et sept mots²¹. L'abrégé d'une demande selon le PCT, qui compte entre 50 et 150 mots²², comprend un résumé de la divulgation de l'invention²³. Les rapports de recherche internationale et les rapports préliminaires internationaux sur la brevetabilité sont des documents établis par les offices des brevets relatifs à la brevetabilité des inventions qui comptent en moyenne entre 50 et 200 mots pour les rapports de recherche internationale et entre 550 et 950 mots pour les rapports préliminaires, selon la langue dans laquelle ils ont été rédigés. Le texte contenu dans ces rapports diffère considérablement de celui des demandes selon le système de La Haye.

20. La Division de la traduction du PCT s'appuie sur les mémoires de traduction du PCT créées pour les 18 combinaisons de langues des 10 langues de publication du PCT²⁴, sur la base de données terminologique WIPO Pearl²⁵ et sur WIPO Translate. Le nombre de demandes selon le PCT par langue de publication variant au fil des ans²⁶, un grand nombre d'entrées sont disponibles dans les mémoires de traduction pour les combinaisons linguistiques allemand, chinois, coréen et japonais vers l'anglais, alors qu'à titre de comparaison, peu d'entrées sont disponibles dans les combinaisons linguistiques arabe et russe vers l'anglais²⁷. En conséquence, les traductions de l'arabe et du russe vers l'anglais sont effectuées par des traducteurs humains (à l'aide des rares mémoires de traduction existantes et de la base de données terminologique).

21. De même, WIPO Translate n'est actuellement utilisé qu'en interne pour traduire du chinois vers l'anglais et de l'anglais vers le français en raison des différents volumes de données d'apprentissage disponibles. L'utilisation de WIPO Translate est actuellement étudiée pour d'autres combinaisons linguistiques (notamment pour le français, l'allemand, le coréen et le japonais vers l'anglais).

22. Le modèle WIPO Translate adapté aux traductions internes du PCT n'est d'aucune utilité pour traduire des documents du système de La Haye et ce, pour deux raisons. Tout d'abord, il n'est actuellement utilisé que dans deux combinaisons linguistiques et ensuite, il a été entraîné avec du contenu tiré d'un domaine technique différent (documents de brevet).

DIVISION LINGUISTIQUE

23. La Division linguistique de l'OMPI est chargée d'assurer des services de traduction de traités et autres instruments internationaux administrés par l'OMPI et l'UPOV²⁸, de textes législatifs nationaux, de documents pour les assemblées des États membres, d'autres organes principaux, de comités et de groupes de travail (notamment leurs documents de travail et

²¹ Voir la règle 4.3 du règlement d'exécution du PCT.

²² Lorsque l'abrégé est établi ou traduit en anglais. Voir la règle 8 du règlement d'exécution du PCT.

²³ Voir la règle 8.1 a) du règlement d'exécution du PCT.

²⁴ Cependant, la Division de la traduction du PCT traduisant vers l'anglais, les entrées dans les autres combinaisons linguistiques, c'est-à-dire de l'anglais vers les autres langues de publication du PCT, sont quasi-inexistantes, exception faite de la combinaison l'anglais vers le français.

²⁵ La base de données terminologiques du PCT contient des termes scientifiques et techniques issus des documents de brevet dans les 10 langues de publication du PCT. Elle compte actuellement quelque 230 000 termes. Elle est accessible au public à l'adresse <https://www.wipo.int/reference/fr/wipopearl>

²⁶ Voir le graphique C1 de la *Revue annuelle du PCT 2022*.

²⁷ De plus, les rapports de recherche internationale et les rapports préliminaires internationaux sur la brevetabilité établis en arabe ou en russe ne sont pas disponibles dans un format exploitable par machine, raison pour laquelle les traductions de ces documents ne peuvent être ajoutées aux mémoires de traduction du PCT ou utilisées pour entraîner WIPO Translate. Par conséquent, les données disponibles dans les mémoires de traduction du PCT pour l'arabe et le russe proviennent uniquement des traductions des titres et des abrégés des demandes selon le PCT déposées dans ces langues.

²⁸ Des services de traduction et d'interprétation sont également fournis à l'UPOV, dont les quatre langues officielles sont le français, l'allemand, l'anglais et l'espagnol.

rapports de réunions), de publications, de matériel de formation, du site Web de l'OMPI et d'autres contenus de communication, de et vers les six langues officielles des Nations Unies. Le texte traduit par la Division linguistique diffère donc du texte contenu dans les demandes selon le système de La Haye²⁹.

24. Les traductions³⁰ de la Division linguistique de l'OMPI sont effectuées pour la plupart de l'anglais vers les autres langues officielles des Nations Unies, quelques-unes sont effectuées du français et de l'espagnol vers l'anglais et une infime partie du contenu est traduite des autres langues officielles des Nations Unies ainsi que de l'allemand et du portugais vers l'anglais, et inversement³¹. Toutes les traductions sont effectuées par des traducteurs humains, à l'aide des mémoires de traduction correspondantes. Il convient de remarquer à cet égard que les mémoires de traduction utilisées par la Division linguistique contiennent des données de l'anglais vers l'arabe, le chinois et le russe mais qu'elle ne dispose d'aucune source de données facilement accessible de ces langues vers l'anglais. Depuis quelques années, les résultats de la postédition de la traduction automatique générée par WIPO Translate ont été utilisés pour traduire certains types de documents, tels que les rapports in extenso.

25. Le modèle WIPO Translate réservé à l'usage interne de la Division linguistique ne serait d'aucune utilité pour traduire des documents du système de La Haye et ce, pour deux raisons. Tout d'abord, il a été entraîné avec du texte tiré d'un domaine technique différent (documents de réunions) et ensuite, la majorité des traductions sont effectuées de l'anglais vers d'autres langues, raison pour laquelle aucune donnée n'est disponible de ces autres langues vers l'anglais.

IV. AUTRES CONSIDÉRATIONS

QUALITÉ DE LA TRADUCTION

26. Il est important de rappeler que la traduction de tout texte contenu dans une demande selon le système de La Haye est établie par le Bureau international sans aucune contribution ni aucun contrôle de la part du déposant³². Plus précisément, l'indication du produit et la description, ou leur traduction, sont ensuite prises en considération par certains offices désignés lors de l'examen quant au fond du dessin ou modèle et conservent une importance pour les procédures administratives ou judiciaires dans toutes les parties contractantes désignées en matière de refus, d'invalidation ou d'atteinte. Par conséquent, les traductions établies par le Bureau international représentent un élément important pour toute décision quant au fond concernant la validité ou l'étendue de la protection dans les ressorts juridiques des parties contractantes³³ alors que, contrairement à ce qui se passe dans le cadre du PCT³⁴, le déposant n'y participe pas. Par conséquent, une traduction de qualité par le Bureau

²⁹ Bien que la Division linguistique soit chargée de la traduction du contenu relatif au système de La Haye, comme les documents du groupe de travail, ce texte diffère de celui des demandes selon le système de La Haye.

³⁰ Environ 90% selon les estimations de la Division linguistique de l'OMPI.

³¹ Les traductions de et vers l'allemand et le portugais sont externalisées.

³² Voir la règle 6.4) du règlement d'exécution commun. À cet égard, il est noté que, bien que le déposant ait la possibilité de joindre à la demande internationale une proposition de traduction du texte contenu dans la demande internationale, le Bureau international n'en reçoit en fait jamais.

³³ Les parties contractantes peuvent se fonder sur la description telle qu'elle a été déposée ou traduite.

³⁴ L'abrégé d'une demande selon le PCT, traduit par le Bureau international, sert simplement d'outil de sélection aux fins de la recherche dans le domaine technique considéré et n'est pas pris en considération à d'autres fins, comme l'étendue de la protection demandée (article 3.3) du PCT et règle 8.3) du règlement d'exécution du PCT). L'étendue de la protection est définie par les revendications, tandis que la description divulgue l'invention. Les revendications et les descriptions sont toutefois traduites par le déposant lors de l'ouverture de la phase nationale devant un office, et l'examen quant au fond par un office national ou régional est effectué sur la base de cette traduction.

international est impérative pour la préservation des droits des titulaires sur le territoire des parties contractantes désignées.

ÉVALUATION DES OUTILS DE TRADUCTION AUTOMATIQUE NEURONALE

27. En outre, le Bureau international a procédé à un test d'évaluation de la facilité d'utilisation des outils de traduction automatique neuronale, notamment WIPO Translate qui est accessible au public sur le site de Patentscope, pour la traduction du texte contenu dans les demandes selon le système de La Haye. Le test d'évaluation a été réalisé pour la traduction de texte contenu dans les enregistrements nationaux chinois de dessins ou modèles, qui ont été extraits de la Base de données mondiale sur les dessins et modèles³⁵. Selon les premières conclusions des traducteurs internes, la qualité du résultat n'était pas suffisante pour une utilisation en postédition, à moins d'avoir entraîné les outils sur des textes propres au domaine. WIPO Translate présentait cependant certains avantages par rapport aux autres outils de traduction automatique neuronale pour la traduction de termes techniques.

28. Une évaluation plus approfondie de la facilité d'utilisation de WIPO Translate pour la traduction des textes de La Haye en chinois et dans d'autres langues pourrait être réalisée. Pour cela, le Bureau international aurait toutefois besoin d'échantillons d'enregistrements de dessins ou modèles dans d'autres langues, couvrant une grande variété d'indications, de légendes et de descriptions de produits³⁶.

SERVICES DE TRADUCTION AUTOMATIQUE FOURNIS PAR DES PRESTATAIRES EXTERNES

29. Une autre possibilité, le cas échéant, consisterait à recourir à des services de traduction automatique fournis par des prestataires externes. Cependant, la qualité de la traduction obtenue ne sera peut-être pas aussi élevée qu'avec WIPO Translate, car elle dépend de l'entraînement avec un grand volume de données dans le domaine concerné. À ce sujet, il est rappelé que le modèle de WIPO Translate pour les traductions relatives au système de La Haye est entraîné avec le corpus des mémoires de traduction de La Haye. De plus, le recours à des services externes de traduction automatique peut également être plus coûteux que WIPO Translate, qui est développé en interne. Enfin, en raison de la nature confidentielle des demandes selon le système de La Haye, il conviendrait de procéder à une évaluation des prestataires potentiels en termes de sécurité et de protection des données.

V. FACILITÉ D'UTILISATION DES OUTILS DE TRADUCTION AUTOMATIQUE POUR L'INTRODUCTION DE LANGUES SUPPLÉMENTAIRES

30. Actuellement, le Bureau international ne dispose pas de mémoires de traduction pour la traduction du texte contenu dans les demandes selon le système de La Haye dans d'autres langues que les trois langues de travail actuelles. De plus, WIPO Translate n'a pas été entraîné pour la terminologie de La Haye dans d'autres langues. À la suite de l'analyse de la disponibilité et de la facilité d'utilisation des technologies de traduction indiquées ci-dessus, la traduction des textes relatifs au système de La Haye à partir de et vers toute nouvelle langue sera, pour les années à venir, confiée à des traducteurs humains.

31. En outre, la facilité d'utilisation de ces outils pour d'autres langues dépendrait de plusieurs facteurs, notamment du nombre de demandes déposées dans cette langue (qui détermine le volume des entrées dans les mémoires de traduction et les données d'entraînement pour WIPO Translate), du volume disponible d'autres données d'entraînement pertinentes dans

³⁵ Le test d'évaluation a concerné la traduction de textes du chinois vers l'anglais, car WIPO Translate est utilisé pour les traductions effectuées par le PCT dans cette combinaison linguistique.

³⁶ À cet égard, il est à noter que la Base de données mondiale sur les dessins et modèles ne contient aucune collection en arabe ou en russe.

chaque combinaison linguistique pour l'entraînement de WIPO Translate et de la similarité ou de la différence relative des différentes combinaisons linguistiques.

CONSIDÉRATIONS RELATIVES À LA CONSTITUTION DE MÉMOIRES DE TRADUCTION

32. Si des langues supplémentaires devaient être introduites dans le système de La Haye, le Bureau international devrait créer et développer des mémoires de traduction pour chaque nouvelle combinaison linguistique.

33. À cet égard, la constitution d'un corpus de mémoires de traduction avant l'introduction d'une langue supplémentaire pourrait permettre de faciliter leur utilisation et réduire les coûts de traduction si un nombre suffisant de dépôts peut être prévu dans les langues en question. La traduction en amont des entrées actuelles des mémoires de traduction dans des langues supplémentaires, qui est actuellement à l'étude par le groupe de travail de Madrid pour les indications de produits et de services, pourrait donc être une option possible. Toutefois, si la création de termes équivalents dans d'autres langues peut s'avérer utile pour des segments de texte plus courts (notamment les indications de produits ou les légendes), elle n'aurait pas la même utilité pour la traduction des descriptions³⁷, qui représentent l'essentiel du travail de traduction.

34. De plus, la traduction des entrées dans les mémoires de traduction aurait un coût³⁸. Par conséquent, il conviendrait de comparer le coût de la traduction des mémoires de traduction existantes dans des langues supplémentaires avec le coût de la traduction dans des langues supplémentaires uniquement du texte contenu dans les demandes selon le système de La Haye et de la constitution de mémoires de traduction au fil du temps. En dessous d'un certain nombre de dépôts, le coût de la traduction des mémoires de traduction existantes serait sensiblement plus élevé que celui de la traduction du texte contenu dans les demandes selon le système de La Haye déposées dans des langues supplémentaires.

ENTRAÎNEMENT DE WIPO TRANSLATE

35. Le Bureau international se sert actuellement de WIPO Translate pour la traduction automatique des textes de La Haye qui ne figurent pas dans les mémoires de traduction. L'utilisation de WIPO Translate pour les nouvelles combinaisons linguistiques serait une option.

36. À cet égard, afin d'obtenir une qualité suffisante pour la traduction des textes de La Haye, il faudrait entraîner WIPO Translate avec de grands volumes de textes dans le domaine concerné pour chaque nouvelle combinaison linguistique. Ainsi, le partage de toute combinaison linguistique pertinente de termes concernant les indications de produits, les légendes, les revendications et les descriptions serait utile. Il n'est cependant pas certain que les offices des parties contractantes eux-mêmes détiennent de grands volumes de ces données d'entraînement ou que de tels volumes soient disponibles ailleurs. Il convient également de souligner que, si WIPO Translate peut être entraîné avec des entrées dans un sens de traduction (par exemple, de l'anglais vers le chinois), cela ne garantit pas la même qualité de traduction s'il est utilisé pour des traductions dans l'autre sens (par exemple, du chinois vers l'anglais)³⁹. En outre, la traduction initiale du texte pertinent aurait un coût de toute façon, si elle devait être effectuée par le Bureau international⁴⁰.

³⁷ Comme indiqué ci-dessus, les descriptions sont rédigées en texte libre et présentent une grande variété. Voir également le paragraphe 17.

³⁸ Les coûts de la traduction des mémoires de traduction relatives au système de Madrid en arabe, chinois et russe sont estimés entre 2,36 et 4,31 millions de francs suisses. Voir le paragraphe 42 du document MM/LD/WG/19/7.

³⁹ Voir la note de bas de page 8.

⁴⁰ Voir le paragraphe 34.

37. En outre, il est rappelé que, bien qu'il existe des volumes considérables de données d'entraînement potentielles, le système du PCT n'utilise actuellement WIPO Translate que pour deux combinaisons linguistiques⁴¹ et que la Division linguistique ne l'utilise pour le moment que dans une mesure très limitée⁴². Par conséquent, à ce stade, l'utilisation potentielle de WIPO Translate pour la traduction des textes de La Haye au-delà des trois langues actuelles semble assez limitée.

38. Enfin, il est également noté que la traduction automatique neuronale, y compris WIPO Translate, progresse et qu'une évaluation régulière des outils de traduction automatique neuronale disponibles pour la traduction des textes de La Haye pourrait donc être recommandée.

VI. INCIDENCES SUR LES ESTIMATIONS DE COÛTS

39. Les coûts récurrents pour la traduction des textes de La Haye estimés dans le document H/LD/WG/10/4 étaient calculés sur la base des tarifs applicables dans le système du PCT pour l'externalisation des traductions⁴³. Les taux sont établis en fonction du volume de texte traduit par combinaison linguistique, du caractère technique du texte traduit et des spécificités de chaque combinaison linguistique dans le cadre du PCT. Par conséquent, le taux effectivement applicable pour chacune des langues supplémentaires serait susceptible d'être modifié dans le contexte de la situation générale du système de La Haye, si elles devaient être introduites. À cet égard, la présente analyse examine les effets potentiels d'atténuation des technologies de traduction actuellement disponibles sur ces coûts de traduction récurrents associés à des langues supplémentaires.

40. Les estimations de coûts figurant dans ce document tenaient déjà compte des gains d'efficacité obtenus grâce à la constitution progressive de mémoires de traduction. Même si le Bureau international traduisait le corpus des mémoires de traduction dans des langues supplémentaires, il est peu probable que cela augmente le taux de correspondance pour la traduction des descriptions, et la traduction initiale aurait un coût⁴⁴. En fait, compte tenu de l'évolution continue dans le domaine des dessins et modèles industriels et du texte connexe contenu dans les demandes selon le système de La Haye, le Bureau international observe une diminution de la croissance du taux de correspondance dans les mémoires de traduction de La Haye et estime donc qu'environ 60% des propositions de traduction générées par un outil de traduction automatique neuronale feront toujours l'objet d'une postédition humaine⁴⁵.

41. Par ailleurs, compte tenu de la présente analyse, les outils de traduction automatique actuellement disponibles ne permettraient pas, dans un avenir proche, d'améliorer la productivité ou l'efficacité des traductions dans d'autres langues du texte contenu dans les demandes selon le système de La Haye. Le coût des traductions et du contrôle de la qualité ne pourrait pas être réduit sans incidence sur la qualité de la traduction. À ce sujet, il est souligné que tout compromis sur la qualité de la traduction porterait atteinte aux droits des utilisateurs⁴⁶. Par conséquent, l'analyse technique ne débouche pas sur une révision des estimations de coûts pour l'introduction de langues supplémentaires dans le système de La Haye.

⁴¹ Voir le paragraphe 21.

⁴² Certaines catégories de documents, tels que les rapports in extenso des réunions de l'OMPI. Voir le paragraphe 24.

⁴³ L'estimation des coûts indiquait le coût de traduction supplémentaire pour l'introduction du chinois et du russe, conformément à la demande du groupe de travail (voir le paragraphe 29 du document H/LD/WG/8/8). Bien que l'on puisse supposer que les mêmes coûts récurrents de traduction pour l'introduction d'autres langues pourraient également être calculés sur la base des tarifs applicables dans le système du PCT, une évaluation supplémentaire pourrait être effectuée, si le groupe de travail le décide.

⁴⁴ Voir les paragraphes 34 et 36.

⁴⁵ Voir le paragraphe 13.

⁴⁶ Voir le paragraphe 26.

42. Enfin, il est noté que l'utilisation des technologies de traduction ne peut avoir des incidences que sur le coût des traductions effectives. Tous les autres coûts, tels que le coût des serveurs permettant de faire face à toute charge de travail supplémentaire en matière de traduction et de contrôle de la qualité, s'appliqueraient indépendamment de l'utilisation des technologies de traduction si des langues supplémentaires étaient introduites dans le système de La Haye.

43. Le groupe de travail est invité à examiner le contenu du présent document et à faire part de ses observations à cet égard.

[Fin du document]