F

OMPI
ORGANISATION MONDIALE
DE LA PROPRIÉTÉ
INTELLECTUELLE

CWS/14/7
ORIGINAL : ANGLAIS
DATE : 23 SEPTEMBRE 2026

Comité des normes de l'OMPI (CWS)

Quatorzième session
Genève, 23 – 27 novembre 2026

ANALYSE DES RESULTATS DE L'ENQUETE SUR LA NUMERISATION DE LA DOCUMENTATION EN MATIERE DE PROPRIETE INTELLECTUELLE

Document établi par le Secrétariat

RESUME

1. Le Bureau international a analysé les réponses à l'enquête portant sur les difficultés rencontrées par les offices lors de la numérisation de leurs documents de propriété industrielle. L'analyse de l'enquête révèle que les principales difficultés rencontrées par les offices résident dans l'extraction de contenus complexes à partir des documents de brevet pendant ce processus de numérisation.

RAPPEL

2. À la treizième session du Comité des normes de l'OMPI (CWS), le Bureau international a présenté un document portant sur les difficultés liées à la numérisation qu'il avait notamment constatées au cours de la période où les administrations chargées de la recherche internationale s'efforçaient de satisfaire aux exigences en matière de documentation minimale prévues par le Traité de coopération en matière de brevets (PCT) (voir le document CWS/13/28).

3. Le CWS a approuvé l'envoi, aux membres du CWS et aux parties contractantes du PCT, d'une enquête portant sur les difficultés rencontrées par les offices dans la numérisation de leurs collections de documents de propriété intellectuelle, au moyen d'une circulaire conjointe du CWS et du PCT (voir le paragraphe 212 du document CWS/13/33).

RESULTATS DU QUESTIONNAIRE

4. En mai 2026, le Secrétariat a élaboré et diffusé le questionnaire par l'intermédiaire d'une circulaire conjointe du CWS et du PCT à l'intention des membres du CWS, des offices récepteurs du PCT, des administrations chargées de la recherche internationale et des

administrations chargées de l'examen préliminaire international, ainsi que des offices désignés ou élus (voir C. CWS 206/C. PCT 1701). L'enquête invitait les destinataires de la circulaire à y répondre entre le 15 mai et le 15 juin 2026.

5. Cinquante-six offices ou organisations ont participé à l'enquête, notamment les offices des États membres suivants : Algérie (DZ), Allemagne (DE), Arabie saoudite (SA), Australie (AU), Azerbaïdjan (AZ), Bahreïn (BH), Belgique (BE), Bosnie-Herzégovine (BA), Brésil (BR), Bulgarie (BG), Canada (CA), Chili (CL), Chine (CN), Costa Rica (CR), Croatie (HR), Danemark (DK), Espagne (ES), Estonie (EE), États-Unis d'Amérique (US), Fédération de Russie (RU), France (FR), Géorgie (GE), Guinée équatoriale (GQ), Hongrie (HU), Inde (IN), Irlande (IE), Israël (IL), Jamaïque (JM), Japon (JP), Lettonie (LV), Lituanie (LT), Mexique (MX), Mozambique (MZ), Norvège (NO), Nouvelle-Zélande (NZ), Oman (OM), Paraguay (PY), Pérou (PE), Philippines (PH), Pologne (PL), République de Corée (KR), République de Moldavie (MD), République tchèque (CZ), République-Unie de Tanzanie (TZ), Roumanie (RO), Royaume-Uni (GB), Serbie (RS), Singapour (SG), Slovaquie (SK), Sri Lanka (LK), Suède (SE) et Turquie (TR); ainsi que les organisations suivantes : Office régional africain de la propriété intellectuelle (AP), Organisation eurasienne des brevets (EA), Office européen des brevets (EP) et Office de l'Union européenne pour la propriété intellectuelle (EM).

6. Les réponses individuelles et une synthèse des résultats de l'enquête ont été publiées dans la partie 7 du Manuel de l'OMPI. Un atelier a été tenu en juillet 2026 afin de présenter les résultats et l'analyse préliminaire réalisée par le Secrétariat. De plus amples informations sur les conclusions de cet atelier figurent dans le document CWS/14/8.

ANALYSE DES RESULTATS DE L'ENQUETE

7. En menant cette enquête, le Bureau international souhaitait mieux cerner les difficultés rencontrées par les offices dans la numérisation de leurs collections de documents de propriété intellectuelle afin de formuler des recommandations sur la manière de les surmonter et, ce faisant, d'aider les offices dans leurs activités de numérisation.

8. Même si cette enquête portait sur les difficultés liées à la numérisation de la documentation relative à la propriété intellectuelle, les résultats montrent que ces difficultés concernaient principalement les documents de brevet, plus des trois quarts des participants de l'enquête ayant signalé un problème à cet égard. Ce chiffre est à comparer aux 7% d'offices en ce qui concerne les marques et aux 4% en ce qui concerne les dessins et modèles industriels.

9. La première question posée aux offices dans le cadre de l'enquête visait à déterminer si la contribution aux bases de données mondiales hébergées par le Bureau international pouvait constituer un facteur de motivation pour ces activités de numérisation. Toutefois, aucune corrélation n'a été observée entre les offices ayant mené des activités de numérisation et ceux contribuant aux bases de données hébergées par le Bureau international. Les résultats sont présentés ci-dessous dans la figure 1.

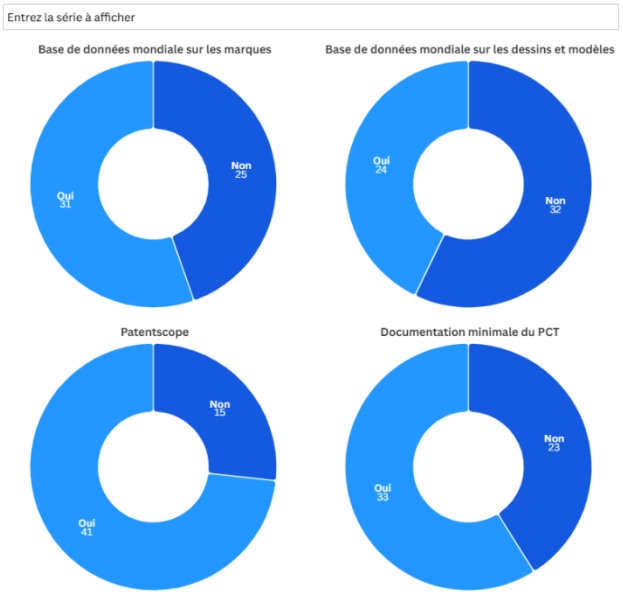


Figure 1 : Pourcentage des participants de l'enquête qui contribuent aux bases de données de l'OMPI.

10. Les questions suivantes portaient sur les formats acceptés par les offices lors du dépôt et sur les formats dans lesquels ceux-ci publient leurs documents. Si ces formats diffèrent, cela signifie que les offices procèdent à une conversion et, dans certains cas, à une numérisation. Les trois principaux formats acceptés lors du dépôt des demandes de brevet étaient le PDF (84%), le papier (82%) et le DOCX (47%). Les trois principaux formats pour les marques et les dessins et modèles étaient le PDF (70 à 80%), les formats fondés sur des images (60 à 70%) et le papier pour les dessins et modèles industriels (environ 70%). Une hypothèse susceptible d'expliquer le nombre élevé d'offices qui acceptent les dépôts sur papier pourrait être qu'il s'agit de parties contractantes du Traité sur le droit des brevets (PLT). Toutefois, ce traité compte 45 parties contractantes, et celles-ci ne correspondent pas aux participants de l'enquête ayant indiqué qu'ils acceptaient les demandes de brevet sur papier.

11. En ce qui concerne les formats utilisés par les offices pour la publication des brevets, le format PDF prédominait largement (plus de 90%), une grande partie des publications étant également conformes aux normes ST.36 ou ST.96 de l'OMPI (plus de 70%). Pour les marques et les dessins et modèles industriels, le format PDF (plus de 80%) et les formats fondés sur des images (plus de 40%) ont été les formats de publication les plus courants, plus de 50% des publications étant conformes à au moins une norme de l'OMPI. Ces résultats indiquent que l'une des raisons pour lesquelles la numérisation pose moins de difficultés pour les marques et les dessins et modèles industriels tient au fait que le format utilisé lors du dépôt est le même que celui utilisé pour la publication.

12. En ce qui concerne les données bibliographiques, les trois principaux formats de publication pour chaque type de droit de propriété intellectuelle examiné sont présentés dans le tableau 1 ci-dessous. Ces résultats indiquent que les offices préfèrent publier leurs données bibliographiques au format PDF ou XML, quel que soit le type de droit de propriété intellectuelle.

Titres de propriété intellectuelle	Format	Nombre d'offices publiant dans le format
Brevets	PDF	43
	XML	34
	Texte brut	14
Marques	PDF	40
	XML	26
	Texte brut	13
Dessins et modèles	PDF	37
	XML	23
	Autre	14

Tableau 1 : Les trois principaux formats de données bibliographiques par type de titre de propriété intellectuelle.

13. La question suivante portait sur la langue de publication. Il s'agissait de déterminer si la langue de dépôt a une incidence sur la facilité avec laquelle les offices peuvent numériser un document de propriété intellectuelle. Les principales langues de dépôt sont présentées dans la figure 2.

14. Aucune difficulté n'a été signalée par les offices, à l'exception de trois d'entre eux qui ont fait état de quelques problèmes mineurs liés à l'alphabet cyrillique. Cela indique que les difficultés liées à la numérisation sont indépendantes de la langue de dépôt.

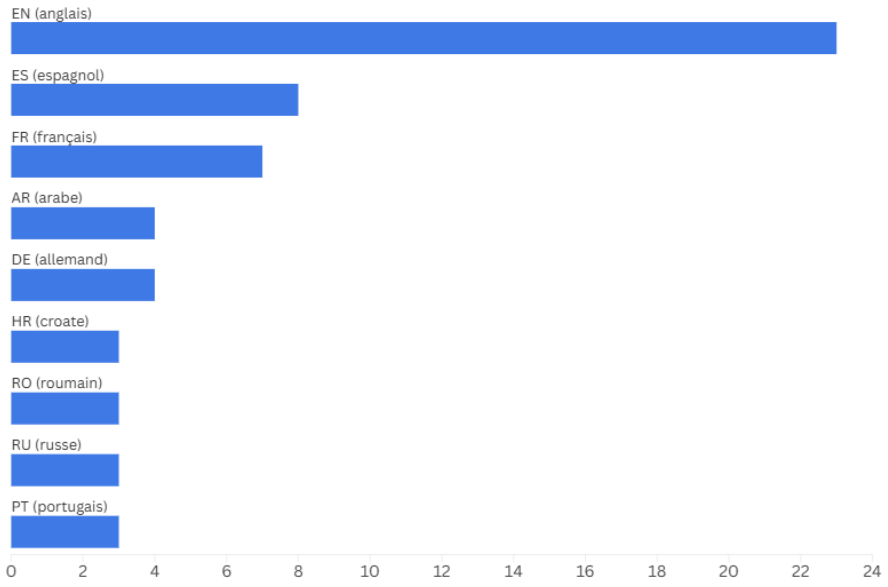


Figure 2 : Principales langues de dépôt parmi les participants de l'enquête

15. Compte tenu des travaux de l'Équipe d'experts chargée du traitement de texte du PCT, la question de savoir si les offices publient les dessins en couleur qu'ils reçoivent ou s'ils se limitent à une publication en noir et blanc présente un intérêt particulier pour le système du PCT. Le Bureau international est déjà en mesure de publier des dessins en couleur si la demande est déposée au format XML, mais il souhaitait mieux cerner les besoins des

utilisateurs avant de proposer des modifications à la règle 11 et à l'annexe F des Instructions administratives du PCT. Les résultats montrent que seule la moitié environ de ces offices publie des dessins en couleur, lorsqu'ils en autorisent le dépôt. Les résultats sont présentés ci-dessous dans la figure 3.

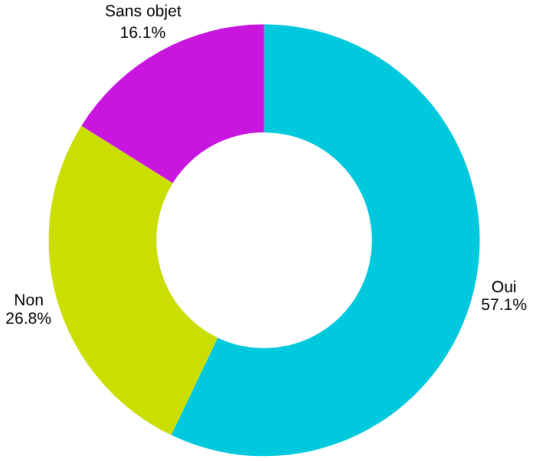


Figure 3 : Les offices publient-ils des dessins en couleur?

16. La dernière partie de l'enquête comportait des questions portant sur les activités de numérisation menées spécifiquement par les offices, ainsi que sur les procédures et les plateformes qu'ils utilisaient. Dans ce contexte, la numérisation était définie comme "le processus consistant à transcrire des données sous une forme numérique pouvant être directement traitée par un ordinateur"[1]. Plus de 80% des participants de l'enquête ont indiqué mener des activités de numérisation. Les autres se contentent de publier les formats tels qu'ils les ont reçus lors du dépôt.

17. En fonction de leurs procédures de numérisation, les offices peuvent être classés dans l'une des cinq catégories suivantes :

- Catégorie 1 : Numérisation de base (sans reconnaissance optique de caractères);
- Catégorie 2 : Numérisation puis reconnaissance optique de caractères pour générer un fichier PDF;
- Catégorie 3 : Numérisation et reconnaissance optique de caractères pour générer un fichier structuré (au format XML, par exemple);
- Catégorie 4 : Comme ci-dessus, mais avec un flux optimisé par l'IA; et
- Catégorie 5 : Des flux de travail entièrement numériques, du dépôt à la publication.

18. Lorsque les offices externalisent leurs activités de numérisation, les principales raisons invoquées sont généralement le volume important de documents à numériser, notamment en ce qui concerne les fonds d'archives de brevets, l'expertise spécialisée requise pour mener à bien cette tâche, la rentabilité de l'externalisation et la rapidité avec laquelle celle-ci peut être réalisée en externe.

19. En réponse à la question initiale et à l'objet principal de l'enquête, les réponses indiquent que les principaux défis rencontrés par les offices concernent l'extraction de contenus complexes à partir de documents de brevet. Même les offices de la catégorie 4 ont indiqué qu'ils continuaient à rencontrer des difficultés, y compris lorsqu'ils utilisaient l'intelligence artificielle pour extraire les contenus, notamment en ce qui concerne l'extraction de formules

[1] Collins English Dictionary

chimiques et d'expressions mathématiques. Parmi les autres défis majeurs figurent la mauvaise qualité des documents papier originaux, les problèmes liés à la structure des documents (par exemple, la numérotation des paragraphes et des revendications) et l'extraction des dessins à partir du texte qui les entoure.

PROCHAINES ETAPES

20. À la lumière des réponses à l'enquête et de l'analyse présentée aux paragraphes 7 à 18 ci-dessus, le Bureau international envisagera l'élaboration d'un ensemble de lignes directrices en matière de numérisation, en formulant, dans la mesure du possible, des recommandations visant à relever les défis mis en évidence par cette enquête, ou en préconisant l'utilisation d'un outil commun tel que la solution ROC de l'OMPI. De plus amples informations concernant ces prochaines étapes figurent dans le document CWS/14/8.

21. En ce qui concerne les projets futurs, 11% des offices ont indiqué dans leurs réponses à l'enquête leur intention de mettre en œuvre les normes de l'OMPI ou d'étendre leur utilisation à l'avenir. Les offices souhaitant mettre en œuvre les normes de l'OMPI sont invités à solliciter une assistance technique en envoyant un courrier électronique au Bureau international à l'adresse cws.mail@wipo.int.

22. Le CWS est invité

a) à prendre note du contenu du présent document,

b) à examiner et à approuver l'analyse des résultats de l'enquête telle qu'elle figure aux paragraphes 7 à 18, et à approuver leur publication dans la partie 7 du Manuel de l'OMPI, et

c) à encourager ses membres à s'adresser au Bureau international pour obtenir une assistance technique s'ils souhaitent mettre en œuvre les résultats structurés dans le cadre de leurs activités de numérisation.

[Fin du document]