

## Comité de Normas Técnicas de la OMPI (CWS)

**Decimotercera sesión**  
**Ginebra, 10 a 14 de noviembre de 2025**

### DIGITALIZACIÓN DE DOCUMENTACIÓN SOBRE PROPIEDAD INTELECTUAL

*Documento preparado por la Oficina Internacional*

#### RESUMEN

1. Los miembros del Comité de Normas Técnicas de la OMPI (CWS) han indicado que uno de sus principales retos es la digitalización de su documentación sobre propiedad intelectual (PI). Al mismo tiempo, las autoridades de patentes que desean incluir sus documentos de patente como parte de la documentación mínima del Tratado de Cooperación en materia de Patentes (PCT) están trabajando actualmente para cumplir los requisitos establecidos en las Instrucciones Administrativas del PCT. En este documento se detallan algunos de los retos a los que se enfrentan las Oficinas y se expone la necesidad de elaborar una serie de directrices de digitalización, centradas inicialmente en las colecciones de patentes.

#### ANTECEDENTES

2. En 2024, la Secretaría realizó una encuesta entre los miembros del CWS para conocer los problemas a los que se enfrentan a la hora de intercambiar datos de propiedad intelectual y las posibles soluciones. Los resultados de la encuesta revelaron que uno de los principales problemas a los que se enfrentan las oficinas, independientemente de su tamaño, es que la información de los boletines de propiedad intelectual no está disponible en un formato legible por máquina y que no se dispone de suficientes recursos, tanto en términos de personal cualificado como de recursos informáticos, para apoyar estas actividades (véase el párrafo 8 del documento CWS/12/23).

3. El Equipo Técnico sobre Documentación Mínima del PCT, dirigido por la Oficina Europea de Patentes (OEP) y la Oficina de Patentes y Marcas de los Estados Unidos (USPTO), se creó en 2005 por la Reunión de las Administraciones Internacionales (MIA) en el marco del PCT para llevar a cabo una revisión exhaustiva de dicha documentación. No obstante, los trabajos

no comenzaron en serio hasta principios de 2017, cuando se definió un plan de trabajo para alcanzar los cuatro objetivos siguientes (véase el Apéndice del documento PCT/MIA/24/4):

- a) Objetivo A: Crear un inventario actualizado de la documentación mínima del PCT, tanto de literatura de patentes como de literatura no relacionada con patentes.
- b) Objetivo B: Recomendar criterios y normas para incluir una colección de patentes nacionales en la documentación mínima del PCT.
- c) Objetivo C: Proponer componentes bibliográficos y textuales claramente definidos de los datos de las patentes que deben estar presentes en las colecciones de patentes pertenecientes a la documentación mínima del PCT.
- d) Objetivo D: Recomendar criterios y normas para la revisión, incorporación y mantenimiento de la bibliografía no relacionada con patentes y los conocimientos tradicionales basados en el estado de la técnica. Posteriormente, evaluar la propuesta revisada de las autoridades indias sobre la base de datos de la Biblioteca Digital de Conocimientos Tradicionales de la India, según los criterios establecidos.

4. Tras los debates mantenidos en el Equipo Técnico y en el Grupo de Trabajo del PCT en julio de 2023, la Asamblea de la Unión del PCT aprobó, con efecto a partir del 1 de enero de 2026, modificaciones de las reglas 34, 36 y 63 del reglamento del PCT (documento PCT/A/55/2). En junio de 2024 se promulgaron modificaciones de las Instrucciones Administrativas del PCT (Circular C. PCT 1672), incluida la adición de un nuevo anexo H. En este anexo se detallan los requisitos técnicos destinados a garantizar que todas las colecciones de patentes y modelos de utilidad pertenecientes a la documentación mínima del PCT estén disponibles de forma gratuita para cada Administración encargada de la búsqueda internacional (ISA), en un formato legible por máquina, al menos para las publicaciones a partir del 1 de enero de 2026. Además, se establece que las publicaciones anteriores al 1 de enero de 1991 deberán estar disponibles en formato de texto completo durante los diez años siguientes.

5. El Anexo H especifica que hay tres formatos admisibles para los datos de búsqueda de texto: Instancias XML conformes a la norma ST.36 o a la norma ST.96 de la OMPI, o en formato de texto sin formato. Las autoridades pueden elegir el formato que consideren más adecuado, siempre que permita realizar búsquedas de alta calidad en sus colecciones de patentes. La Oficina Internacional recomienda el uso de XML, al menos para proporcionar datos bibliográficos. Se pide a las oficinas que elijan el texto sin formato como formato preferido y que consideren la posibilidad de migrar a una de las normas técnicas de la OMPI, con el fin de mejorar la calidad de sus publicaciones de patentes y su capacidad de búsqueda.

## PREPARACIÓN

6. Para ayudar a las ISA a cumplir los requisitos técnicos antes de la fecha límite del 1 de enero de 2026, la Oficina Internacional organizó una serie de consultas bilaterales sobre la documentación mínima del PCT en 2025. Las oficinas prepararon con antelación las preguntas que tenían para la Oficina Internacional, y esta se aseguró de que el experto adecuado estuviera presente en la consulta para responderlas. Durante estas sesiones, se observó que las Instrucciones Administrativas del PCT no ofrecían orientaciones sobre la calidad de los documentos de texto completo. A petición suya, la Oficina Internacional compartió con el Equipo Técnico del Fichero de Referencia del CWS su idea de un conjunto de requisitos mínimos para lo que ellos consideran un documento legible por máquina “suficientemente bueno”.

7. En el caso de las Oficinas que están digitalizando sus documentos de patente, a continuación se ofrecen orientaciones para ajustarse a los requisitos del Anexo H:

- En el caso de las instancias XML conformes con la norma ST.36 de la OMPI, deben estar bien formadas y contener un elemento raíz coherente y con el mínimo de errores de reconocimiento óptico de caracteres. Los datos bibliográficos deben incluir, al menos, la referencia de la solicitud, la referencia de la publicación y el título de la invención.
- Las instancias XML que cumplan con la norma WIPO ST.96 deben cumplir con el componente PatentPublication a nivel de documento de dicha norma, pero no es necesario implementar MathML ni OASISTable.
- En el caso de los archivos de texto sin formato, los errores de OCR deben ser mínimos; el texto debe presentarse en párrafos; no deben incluirse números de línea ni de página; y debe permitir un orden de lectura correcto. Deben ir acompañados preferiblemente de datos bibliográficos en formato XML.

8. El Equipo Técnico sobre Documentación Mínima del PCT acordó un proceso de certificación en tres fases para las ISA u otras Oficinas de patentes que deseen que sus colecciones formen parte de la documentación mínima del PCT. En la segunda fase, denominada “fase de verificación”, se comprobó si una oficina puede proporcionar un depósito de datos que contenga documentos conformes a uno de los tres formatos indicados en el anexo H para su descarga y si puede producir un fichero de referencia conforme a la Norma ST.37 de la OMPI, incluidos los indicadores necesarios para la búsqueda de texto.

#### RETOS A LOS QUE SE ENFRENTAN LAS OFICINAS

9. Durante las consultas sobre la documentación mínima del PCT y la fase de verificación, la Oficina Internacional proporcionó comentarios para contribuir a la elaboración de publicaciones de patentes legibles por máquina y a la producción de ficheros de referencia conformes con la Norma ST.37 de la OMPI. Tras este proceso de revisión, las oficinas se enfrentaron a algunos de los retos más habituales a la hora de crear documentos legibles por máquina:

- a) En todos los formatos, la baja calidad del reconocimiento óptico de caracteres (OCR) provoca numerosos errores que contaminan los índices de búsqueda, sobre todo en el caso de los documentos de patentes con fórmulas químicas.
- b) En todos los formatos, se produce una mala detección o tratamiento de la maquetación, lo que provoca la inclusión de información de encabezado o pie de página dentro del texto o la aparición de números de línea.
- c) En todos los formatos, la segmentación de las secciones de descripción, reivindicaciones y dibujos es deficiente, lo que provoca, por ejemplo, que el texto de la descripción se incluya en el texto completo de las reivindicaciones.
- d) En todos los formatos, el lenguaje de las secciones de texto completo está mal etiquetado, lo que provoca problemas críticos de indexación en los sistemas de búsqueda.
- e) En el caso de las instancias conformes con la norma ST.96 de la OMPI, se trata de instancias inválidas que no han sido validadas con respecto al componente de nivel de documento PatentPublication y de la confusión en torno al uso de los códigos de idioma dentro de la jerarquía.
- f) En el caso de las instancias conformes con la norma ST.36 de la OMPI, existe un malentendido sobre lo que se considera una instancia “válida” y cuál es el nivel mínimo

para considerar que el documento XML es “suficientemente bueno”. Por ejemplo, si hay que utilizar el modelo DTD de ST.36.

g) En el caso del formato de texto sin formato, el texto no se proporciona como párrafos propiamente dichos, sino como líneas, lo que dificulta la búsqueda de contenidos. Además, los datos bibliográficos se proporcionan en formatos que no son los de las normas de la OMPI.

## ASISTENCIA TÉCNICA

10. Para hacer frente a algunos de los retos planteados anteriormente, la Oficina Internacional presta asistencia técnica a las oficinas que la solicitan. La solución OCR de la OMPI permite generar instancias XML compatibles con la norma ST.36 de la OMPI a partir de archivos PDF con marcadores basados en imágenes y con función de búsqueda de texto. No obstante, el CWS debe tener en cuenta que esta herramienta no permite detectar ni etiquetar números de reivindicación y trata los cuadros y las fórmulas químicas y matemáticas complejas como dibujos integrados.

11. *OMPI Publish* permite a las Oficinas extraer y controlar la calidad de sus datos en el formato de la Norma ST.36 y, en breve, la salida migrará para ajustarse a la Norma ST.96. Esto facilita el intercambio de información del boletín de muchas Oficinas de PI. En 2025, la Oficina Internacional añadió funciones para extraer documentos de patente para las Oficinas que aplican los requisitos mínimos de documentación del PCT. Se ha completado con éxito un proyecto piloto con la Oficina de Propiedad Intelectual de Filipinas. El mismo enfoque podría aplicarse a otras Oficinas de patentes cuyas colecciones forman parte de la documentación mínima del PCT; las Oficinas interesadas podrían ponerse en contacto con la Oficina Internacional.

12. La Oficina Internacional también presta apoyo directo a unas treinta oficinas para la digitalización de su documentación, incluido el reconocimiento óptico de caracteres (OCR) del texto completo para algunas de ellas. Este proceso consiste en escanear los documentos, digitalizarlos y comprobar la calidad de los datos bibliográficos. En todos los casos, este trabajo se subcontrata a terceros.

## PROPUESTA DE PRÓXIMOS PASOS

13. Teniendo en cuenta estos retos, la Oficina Internacional propone elaborar un conjunto de directrices para la digitalización de documentos de PI con el fin de paliar la falta de comprensión señalada anteriormente. Con el fin de seguir examinando las dificultades que experimentan las oficinas que actualmente digitalizan su documentación de PI, la Oficina Internacional propone realizar una encuesta mediante la publicación de una circular conjunta del CWS y el PCT. La Oficina Internacional tiene previsto preparar esta encuesta y enviar una invitación a los miembros de ambos organismos para que la rellenen durante el primer trimestre de 2026.

14. Aunque en esta fase el debate se centra en la digitalización de colecciones de patentes, es probable que el alcance de estas directrices se amplíe para incorporar nuevas orientaciones sobre la digitalización de documentos de marcas y diseños industriales.

15. Una vez finalizado el plazo de la encuesta, la Oficina Internacional tiene previsto organizar una reunión para debatir sobre la elaboración de las directrices de digitalización en el primer semestre de 2026. Se invitará a todos los miembros del CWS y del PCT. No obstante, los detalles de la reunión, incluido el formato y las fechas, se anunciarán oportunamente. En esta reunión, la Oficina Internacional tiene previsto informar sobre los resultados de la encuesta y sobre un posible calendario para elaborar las directrices de digitalización de los documentos de PI. Se anima a los miembros de CWS a participar en la encuesta y en esta reunión, ya que sus opiniones son fundamentales para elaborar las directrices de digitalización. La Oficina

Internacional también tiene la intención de informar sobre los resultados de esta reunión en la decimocuarta sesión del CWS.

16. Además, la Oficina Internacional tiene previsto organizar un taller sobre las Normas ST.36 y ST.96 de la OMPI en el segundo trimestre de 2026. Se enviarán invitaciones a este taller a los miembros y observadores del CWS y se anunciará el evento en el sitio web de la OMPI. La Oficina Internacional también invita a las oficinas que aplican las normas técnicas de la OMPI a solicitar asistencia técnica cuando la necesiten.

17. Se *invita al CWS a:*

- a) *tomar nota del contenido del presente documento, en particular de los retos a los que se enfrentan las Oficinas, como se indica en el párrafo 9;*
- b) *aprobar la propuesta de realización de la encuesta a través de la circular conjunta en la que se invita a los miembros del CWS y del PCT, tal como se indica en el párrafo 13;*
- c) *alentar a sus miembros a participar en la encuesta y en la reunión, y tomar nota de que los resultados de esta se presentarán en la decimocuarta sesión del CWS, tal y como se indica en el párrafo 15; y*
- d) *alentar a sus miembros y observadores a participar en el taller sobre las Normas ST.36 y ST.96 de la OMPI que se celebrará en el segundo trimestre de 2026, tal y como se indica en el párrafo 16.*

[Fin del documento]