

专利合作条约（PCT）工作组

第十四届会议

2021 年 6 月 14 日至 17 日，日内瓦

以全文格式处理国际申请

国际局编拟的文件

概 述

1. 本文件提出了转为全文处理国际申请的提案。国际局希望修改行政规程附件 F，以便在国际阶段处理由申请人提交或由主管局处理的 XML 全文格式的申请体文件，而这又可以推动向全文国际公布的过渡。为了鼓励更多地提交全文申请，国际局将通过网络服务向所有知识产权局提供它的 DOCX 转换器和全文比对工具，这将作为处理以 DOCX 格式提交的申请的参考实现，并确保从 DOCX 转换为 XML 过程的一致性。

背 景

2. 近年来，PCT 体系中国际申请的提交和处理已从纸质形式过渡到电子形式。纸质申请稳步减少，现在电子申请占国际申请的 98% 以上。这一过渡为主管局提供了新的机会，并使申请人获得在线服务成为可能。它还提高了主管局在国际阶段对申请采取必要行动的效率和时效性。然而，尽管处理工作基本上以电子方式进行，但许多主管局以处理和传送纸页的方式处理和传输图像形式的数据，而不使用国际局和履行不同职能的国家局可直接使用的可读文本。此外，在以全文提交申请的情况下，申请的页面图像通常由主管局处理。

3. 在 2018 年 6 月工作组第十一届会议上讨论的总干事备忘录“PCT 体系——概述及未来可能的方向和优先事项”中，列明了 PCT 体系未来发展的四个主要工作领域（见文件 PCT/WG/11/5），技术（信息技术）环境是其中之一。它强调了 PCT 申请数据必须在申请人、受理局、国际单位和国际局之间以

一致和有效的格式传送，然后再传送给指定局。工作组在 2019 年 6 月第十二届会议上讨论了国际局关于 PCT 在线服务的一份文件（见文件 PCT/WG/12/10）。该文件包括在提交和处理国际申请时为了以可读文本格式传送数据而需使用 XML 的最新情况（见文件第 14 至 22 段）。会议报告（文件 PCT/WG/12/25）第 17 至 38 段提供了这些讨论的详细信息。

4. 本文件为转为全文处理国际申请提出了建议。这些建议包括由国际局和各主管局采取协调行动，确保在处理以全文格式提交或处理的国际申请时的一致性。这些建议涉及提交的申请和后续对申请作出的修改，但同样适用于为国际公布和国际检索及初步审查而提供的译文。

全文处理的好处

5. 以全文格式处理申请体文件相比交换图像文件有许多好处，例如：

- 公布后对国际申请的文本检索更加容易和准确；
- 提高国际申请处理的总体效率；
- 搜索引擎可包括公布的申请原文，避免了用光学字符识别（OCR）来生成文件的全文版本，从而消除了光学字符识别过程中可能产生的错误；
- 专利申请的预分类可通过从文本中提取技术术语自动完成；
- 可以为搜索引擎提供所提交申请的机器翻译译文，必要时，可以将文件部分更高效地提供给翻译辅助软件，以协助人工翻译；
- 可以自动检查申请中的形式缺陷，不再需要检查图像复制件，如页边距和字体大小，从而减少国际公布前的任务数量；
- 在处理过程中对申请的修正、改正和更正可以用粒度标记来显示；
- 全文格式可支持与申请一起提交的彩色附图；以及
- 指定局和选定局可提取国际公布的全文，用于国家阶段的处理。

6. 为了充分受益于这些好处，申请必须以全文格式提交，并且所有主管局和国际局对全文统一进行端到端处理。

现状

以全文格式提交国际申请

7. 2020 年，98.3%的国际申请以电子方式提交，而 2015 年全电子申请的比例为 93.6%。然而，按百分比计算，这一时期的 XML 申请基本保持不变：2020 年为 27.9%，2015 年为 28.2%。因此，提高以 XML 提交的国际申请的百分比对过渡到全文处理至关重要。

受理局	全文 XML 申请	PDF 申请	纸质申请	申请总数
JP	49,005	0	308	49,313
KR	19,525	0	150	19,675
CN	7,651	64,336	354	72,341
IB	293	12,963	176	13,432

受理局	全文 XML 申请	PDF 申请	纸质申请	申请总数
EP	138	37,929	806	38,873
US	0	55,665	108	55,773
其他受理局	160	22,341	2,796	25,297
总计	76,772	193,234	4,698	274,704

表 1：按受理局开列的 2020 年国际申请提交格式

8. 表 1 显示，日本特许厅（JPO）和韩国特许厅（KIPO）收到的以 XML 提交的国际申请约占 89%；XML 国际申请在收到的所有国际申请中的占比在这两个主管局分别是 99.4% 和 99.2%。相比之下，虽然国际局作为受理局允许提交 DOCX 文件并根据产权组织标准 ST. 36 转换为 XML 格式，但选择这么做的申请人很少。大多数申请人似乎仍然担心处理全文申请存在风险，因此仍然选择提交 PDF 文件，尽管以 XML 提交可以额外减费 100 瑞士法郎。

9. 其他一些知识产权局正在努力增加 XML 申请提交量。中国国家知识产权局（国知局）和欧洲专利局（欧专局）都在开展项目来为国际申请的全文提交和处理提供便利。关于国家和地区系统，欧专局、法国国家工业产权局（INPI）和美国专利商标局（美国专商局）正在实施处理全文申请的系统。

目前对 XML 格式国际申请的处理

10. 对于以全文格式提交的申请，大部分后续处理是基于根据 XML 制作的图像文件，包括在主管局之间传送图像文件。修正、改正和更正作为替换页处理，失去了以全文格式处理国际申请的好处。

提 案

11. 国际局认为，提高国际申请全文提交和处理的方法是实施一种机制，使申请人能够使用 DOCX 提交申请，并相信申请在转换为 XML 和为国际公布制作页面时将得到一致的处理。行政规程附件 F 允许在转换 DOCX 格式文件的基础上以 XML 格式提交申请体文件，而 DOCX 文件作为转换前的参考文件予以保留。XML 文件构成国际申请的合法副本，申请人在提交前要审查根据 XML 制作的校样副本。

12. 国际局将通过网络服务向所有知识产权局提供其 DOCX 转换器和比对工具。这些工具将作为处理以全文格式提交的申请的参考实现。使用自己工具的主管局将被要求与这些工具的参考实现相匹配，以确保从 DOCX 转换为 XML 过程的一致性。

13. 除了更新 DTD 以实现全文内容的传送外，国际局还建议更新附件 F，以要求在国际阶段对以全文提交的申请以 XML 格式处理。对 DTD 的修改也将支持使用符合通用申请格式（CAF）的标记来标记对申请的修改，并将提供支持处理这些修改的结构。

拟议的修改对国际申请处理工作的影响

提交国际申请

14. 要以全文格式提交，申请人需要提交 DOCX 或 XML 文件。申请软件将提供一个工具，将 DOCX 文件转换为 XML 文件，并重新制作一份校样副本供申请人在提交前审查。提交的申请文件包将包含作为转换前参考的 DOCX 文件，作为转换软件出错时的备份，同时包含用于处理全文申请的 XML 文件。受理局将使用相同的软件来审查国际申请，并将登记本作为 XML 全文文件转发给国际局。

15. 关于附图和申请的其他部分，例如无法以文本格式处理的数学公式和化学结构式，相关信息将以嵌入式图像的形式包含在 XML 申请文件包中。国际局的转换工具会自动处理这些内容。图像被自动转换为黑白 TIFF 图像，尽管开发处理和公布系统的一个关键目标是允许使用彩色图像。

对国际申请的修改

16. 如果申请人需要提交后续文件以对以 XML 格式提交的国际申请进行修改，其过程将与最初提交申请类似。由于修正、改正和更正不能直接插入申请体文件的全文，因此申请人将编辑申请的 DOCX 或 XML 源文件，该文件将（酌情）被转换为 XML 以进行比对。ePCT 中提供的比对工具或受理局提供的类似解决方案将被用于对修改进行标记，将文件转换成 XML 格式的标记申请体文件，并生成修改请求。在向有关主管局提交修改之前，申请人将能够审查对全文申请作出的修改。

17. 在最初实施阶段，预计会要求主管局向国际局提供没有变更标记的 XML 申请体文件的经修订版本。然后，国际局将使用比对工具添加修改标记。接下来，随着主管局逐渐适应实施工作，主管局和国际局将使用与申请人相同的工具对修改进行标记，并在申请体文件中说明它们的状态。全文申请体文件将包括申请的最新版本，并有一个单独的部分来保存修改历史，以便在需要时查看早期版本和修改内容。审查员在考虑拟议的修改时，可以参考一个自动生成的视图，上面显示了拟议修改的申请、差别文件和要求随修改一起提交的说明，如细则 26.4 规定的改正或细则 66.8 规定的修正。

国际公布

18. 对于以全文格式提交或处理的申请，国际局将会全文公布，XML 文件构成合法副本。国际局暂时还将以 TIFF 格式公布这些申请，将其作为页面图像文件，格式与以纸质方式或电子方式 PDF 格式提交的申请的格式通用。

对于申请人的问题

19. 如果程序设计得足够好，提交全文文件供专利申请和公布程序使用对申请人应该是有利的。许多国家局对其公布的申请和授权专利进行排版，因此需要将纸质或 PDF 格式的申请体文件转换为全文格式，这就需要申请人在公布前进行仔细检查。此外，排版后的文件总是与原始文件的布局不同，其格式选择比大多数文字处理程序更为有限。因此，PDF 文件中任何依赖特定布局选项的信息都可能在创建经过排版的公布时丢失，并且可能难以复制。从一开始就收到全文文件并对这些全文版本作出检索和审查对于主管局来说应该更为便利。就申请人而言，这应该在一定程度上解决了展示或转换方面的问题，如果出现问题，还能增加在早期阶段发现和解决问题的机会。

20. 但是，许多 PCT 缔约国的专利律师对这方面的建议一直持谨慎态度。此前转为全文处理的尝试只在其他可用的提交选项对申请人来说显然更加繁琐的国家获得成功。这在国家体系中是实用的，或者在申请人通常使用同一主管局作为受理局和国际检索单位以及（相关时）初步审查单位的情况下是实用的。但是，如果涉及不同的主管局，就需要进行协调，确保一致，以使申请人和主管局都受益。而这将需要律师在这个过程中提供帮助，以确保他们的关键需求得到满足，并适应新的工作方式，如果实施得当，这一工作方式将使他们的客户和整个体系都受益。

21. 一个长期存在的关切是关于从对文本和图形布局具有控制的文字处理格式转换到支持文件公布和交换程序的简化的 XML 格式时可能出现的错误。无论采用哪种系统，都必须确保：

- (a) 将申请体文件从基于页面的布局转换为基于文本的 XML 内容并适当参考图像，这一过程非常可靠，不会影响正在转换的说明书、权利要求或附图的实质意思；以及

(b) 在极少发生的出现转换错误的情况下，检索和审查过程有助于在早期阶段发现这些错误，并且申请人能够在发现错误时予以改正。

22. 此外，对于国际处理来说，转换的一致性至关重要，无论文件提交给哪个主管局，都能得到相同的结果，以便在一个主管局提交的初始申请可以与提交给另一主管局的修正、改正和更正适当地合并和比对。总体来说，国家局之间转换结果的一致性也更有利于降低成本，避免因需要了解对国家和国际处理的不同要求而产生混淆。

23. 考虑到对申请或授权专利排版时只能使用特定而有限的版式设计这一问题，撰写国际申请时应该已经涉及确保该申请不使用时在公布中无法得到可靠复制的版式设计来传达信息。全文提交解决方案本质上是同一个问题。一个设计合理的转换器将保留内容的所有实质意思，前提是原始文件使用 Unicode 字体，并仅使用转换后的格式允许的版式设计选项来披露发明。

24. 仍有待解决的问题——特别是表格的有效呈现，申请人经常使用版式设计“小技巧”，如字体大小、手动换行、文本对齐和手动调整列的大小，以达到特定的视觉安排，而这往往无法在转换后的表格中再现。人们还经常对符号的使用表示关切，这些符号在文字处理器中的显示方式可能由软件决定，申请人无法控制。但是，此类问题可以被发现并已被考虑到，前提是申请人直接选择的字体中的其他字符符合 Unicode 标准。同样，对于产权组织标准 ST. 36 不支持的可编辑（非图像）格式的数学公式等项目，系统也能发现，必要时可以提取图像版本以纳入转换后的文件中。

25. 全彩附图支持已经是公布系统相关开发工作的一个关键目标。为了成功地转为全文处理，在可能难以在产权组织标准 ST. 36 目前提供的选项中充分呈现或转换的披露内容方面，更多来自律师的意见建议十分重要。要解决此类问题，可以通过扩大 XML 标准中直接提供的选项（如对数学和化学公式标准提供更全面的支持），或者（像彩色附图那样）通过增加允许作为参考文件纳入的格式类型。

所有国际申请向全文处理过渡

26. 从长远来看，国际局希望所有国际申请都向全文处理和公布过渡。在这一过渡之后，以纸质提交或以图像格式提交的所有申请都将被转换为 XML，转换后的 XML 文件为合法副本（同样有可能返回到纸质或图像原件以改正错误）。当比如说 98% 以上的国际申请是以全文提交时，这个目标有可能实现。按照目前的申请水平，这需要对大约 7,000 件申请的页面图像进行转换。达到这一门槛后，国际局将在接下来至少五年内继续制作国际申请的 TIFF 图像文件，并对长期需求进行评估。

实施情况

27. 尽管国际局和其他知识产权局将为申请人提交全文国际申请提供工具，但实施本文件中的建议将需要国际局和各主管局进行合作，支持申请人向处理文本而不是页面图像过渡，提供培训，并在需要时提供帮助。在上文第 13 段中的建议生效后，国际局还将根据要求提供指导，特别是向受理局提供处理全文申请方面的指导。

工作组的审议

28. 请工作组审议上文第 13 段中的提案，以及上文第 14 至 27 段中对国际申请的提交和处理所作的相应修改。尤其请工作组就以下软件和程序开发原则达成一致意见：

- 在国际阶段，如果国际申请以全文格式提交，PCT 主管局和单位将以全文而不是页面图像进行处理；

- PCT 主管局和单位将使用申请的 XML 副本作为正式副本来处理全文申请——但 DOCX 文件（或可能支持的任何同等格式）将被视为显示所提交申请的正式视图，以便允许改正在国际阶段（如果可能，也在国家阶段）发现的任何转换错误。
- PCT 主管局和单位应使用国际局提供的参考 DOCX 转换器或与国际局转换器兼容的其他转换器，以便准确识别和标记所提交版本与后来提交的任何修正、改正和更正之间的差异。
- 国际局暂时将以全文和图像格式公布全文申请，并以图像格式公布其他申请；以及
- 当以全文格式提交的国际申请达到足够大的比例时，国际局将最终转为全文处理和公布所有申请。

29. 请工作组确定在开发将文字处理器文件格式转换为简化的 XML 格式的工具时应解决的问题，以及使用和比对经过上述转换的申请体文件的程序。

30. 请工作组就本文件中的提案发表评论意见。

[文件完]