

# 欧洲专利全文文本数据

## (一) 数据编码

EP-TXTO-10-AB

## (二) 数据名称

欧洲专利全文文本数据

## (三) 文献语种

英文、法文、德文

## (四) 更新周期

每周更新

## (五) 数据内容及格式

**数据内容：**包括著录项目、摘要、权利要求、说明书、说明书附图。

**数据格式：**

数据种类	数据格式
文本数据	XML
图形数据	TIF, PDF

## (六) 文件结构

### 1. 概述

本数据的文件结构包括三层文件夹，实体数据存储在三层文件夹下。

第一层文件夹：数据编码与数据名称。数据编码由国别代码（两位大写字母）+文献内容代码（三位大写字母）+文献内容属性代码（一位大写字母）+专利类型代码+文献种类代码组成；数据名称为数据中文全称，例如：CN-BIBO-10-A\_中国发明专利申请公布著录项目数据。

第二层文件夹：数据发布年代，4位数字，YYYY（4位年），例如2014。

第三层文件夹：数据发布期号，为三位流水号，每年从001开始。

### 2. 相关代码说明

**文献内容代码说明：**

**BIB：** 著录项目数据

**ABS：** 摘要数据

**TXT：** 全文文本

**IMG：** 全文图像

**文献内容属性代码说明：**

**O：** 表示原始数据

**T：** 表示翻译数据

**专利类型代码说明：**

10： 发明专利

20： 实用新型

30： 外观设计

**文献种类代码说明：**

A 发明专利申请

B 发明专利授权

U 实用新型专利授权（韩国为实用新型专利申请）

Y 实用新型专利授权（日本 2000 年以前，中国 1993-2010 年 4 月及韩国为实用新型专利授权）

S 外观设计专利授权

**（七）数据说明**

**1.XEPA 压缩包数据路径**

例如：

XEPA2014008.zip\XEPB2014008\DOC\EPNWA1\000\000\002\799\965\DOC0001.xml

路径说明：

XEP+文献类型（1 位大写字母）+4 位公布年+3 位期号.zip

\XEP+文献类型（1 位大写字母）+4 位公布年+3 位期号

\国别代码+NW 或 W1+文献类型

\补 3 位 0

\补 3 位 0

\补 2 位 0+公开号第 1 位数字

\公开号第 2-4 位数字

\公开号第 5-7 位数字

\DOC00001.xml

最底层文件夹内含有：

- 1 个代码化文件“DOC00001.xml”
- 1 个全文说明书图像“DOCUMENT.PDF”
- 摘要附图[0 个或 1 个“imgafNNN.tif”]
- 切图[0 个或多个“imgbNNNN.tif”]
- 说明书附图[0 个或多个“imgfNNNN.tif”]
- 检索报告页[0 个或多个“srepNNNN.tif”]

## 2.XEPB 压缩包数据路径

除文献类型改变外其他路径信息与 XEPA 压缩包一致。

## 3.DTDS 文件夹：存放数据标准 DTD 文件

## 4.主要数据元素

元素名称		元素内容
ep-patent-document		EP 专利信息
AT T	id	id 号
	file	实例 XML 数据文件的文件名
	lang	语言
	country	国家
	doc-number	公布（公告）号
	kind	文献种类代码
	correction-code	修正编码
	date-publ	公布（公告）日
	status	数据更新状态
	sequence	顺序
	dtd-version	dtd 版本
	date-produced	数据生成日期
file-reference-id	文件引用 id	
	doc-page	文件页
	bibliographic-data	著录项目数据
	SDOBI	著录项目数据

		B000	和专利局有关的系统/文件信息
		B100	公开信息
		B110	专利文档号（通常是公布号）
		B130	公开类型
		B190	公布的国家或者组织
		B140	公布日
		B120	某类文档的简称
		B131EP	扩展的文档类型码
		B132EP	使用的原始类型代码
		B133EP	新类型代码
		B150	专利修正信息
		B195	原始设计记录
		B200	专利申请的相关数据
		B210	申请号
		B220	申请日
		B210EP	EPO 的非标准化形式的申请号
		B211EP	EPOQUE 格式的申请号
		B225	接收局信息数据
		B230	其它日期信息
		B240	工业产权生效日
		B250	原始申请使用的语言
		B250EP	允许的非 EPO 工作语言
		B251EP	流程中使用的语言
		B260	申请公布使用的语言
		B270	源设计记录代码
		B300	外国优先权数据
		B310	优先权号
		B310EP	EPO 非标准形式的优先权号
		B311EP	EPOQUE 系统优先权号
		B320	优先权申请日
		B330	规定的国家和组织
		B340	巴黎公约成员国的代码
		B345	同族专利信息
		B345EP	INPADOC 同族专利信息
		B400	公众可访问日期
		B405	专利公告/公报信息
		B410	未审查、未出版的未授权文档
		B420	已经审查、未出版的未授权

				文档
			B430	未审查、已经出版的未授权文档
			B440	已经审查、已经出版的未授权文档
			B450	出版时授权的文档(如美国专利)
			B451EP	宣告日(EPO)
			B452EP	拟授权日
			B453EP	公开日期和相关限制性决定日期
			B460	专利文献权利要求公开日;
			B470	未出版的授权文档
			B472	授权有关的条目
			B476	无效专利
			B477	作为更正出版的文档
			B480	已更正专利文献的发布日
		B500		技术信息
			B510	国际分类号(IPC)数据
			B510EP	第8版IPC分类信息 (2006年后使用的新IPC分类号)
			B520	国内或者国家分类号 (EPO不使用此元素)
			B530	通用十进制分类
			B540	发明创造名称
			B550	关键词或描述字符串
			B560	引文
			B570	摘要或者权利要求
			B580	检索领域
			B590	说明书和附图相关的信息
			B599EP	后续递交的技术数据(EPO)
		B600		其他涉及专利法律和流程方面的国内专利文档
			B610	在先文档的补充
			B620	在先申请分案出来的
			B620EP	其他类型的关系
			B630	继续文档
			B640	被再版的文档
			B645	本标签标识的文档的复审
			B650	涉及同一申请先前公布的文档
			B655	由其他国家/组织先前公布

				的文档
			B660	本标签标识的文档的替换
			B665	专利更正信息
			B670	实用新型所基于的文档
		B700		和文档相关的当事人
			B710	申请人的有关信息
			B720	发明人的有关信息
			B730	委托人相关信息;
			B740	律师、代理人、代表人的相关信息
			B745	处理本文档的人员
			B780	异议有关的数据
			B790	许可有关的数据
		B800		除了巴黎公约以外的国际条约有关的数据
			B810	PCT 指定国家
			B820	PCT 选择国家
			B840	指定的缔约国
			B844EP	申请/专利延伸的国家 (EPO)
			B830	微生物保藏信息
			B850	按照 PCT 第 23 条(1)或第 40 条(1)进入国家阶段的日期
			B860	PCT 或地区申请信息
			B870	PCT 或地区公布信息
			B880	延期公布的检索报告
			B890	经济互助委员会(CEMA)条约有关的数据
		B900		其他数据
			B910	PCT 申请不再有效日
		abstract		摘要
		description		说明书
		invention-title		发明名称
		technical-field		技术领域
		background-art		背景技术
		disclosure		披露
		summary-of-invention		发明概述
		description-of-drawings		说明书附图
		description-of-embodiments		具体化描述
		best-mode		最佳实施例
		mode-for-invention		发明的应用模式或应用方

			法
		industrial-applicability	工业应用
		reference-signs-list	参考符号列表
		reference-to-deposited-biological-material	生物材料提取参考信息
		sequence-list-text	列表顺序文本
		citation-list	引文信息列表
		heading	标题
		claims	权利要求信息
		amended-claims	修改权利要求
		amended-claims-statement	修改权利要求声明
		drawings	附图信息
		ep-reference-list	引用列表