中国专利引文数据数据

（一）数据名称

中国专利引文数据

（二）文献语种

中文

（三）更新周期

周更新

（四）数据内容及格式

1、基本内容

申请信息、公开信息、专利简单引文信息、非专利简单引文信息。

2、特色内容

专利丰富引文信息，非专利丰富引文信息，被引证信息

3、数据格式

XML

（五）文件存储层级结构

1. 概述

本数据数据结构包括两层文件夹，实体数据存储在第二层文件夹下。

第一层文件夹：数据类别。数据类别由数据内容描述表示。例如：enriched citation。

第二层文件夹：发文日期范围，两个8位日期组成时间段，YYYYMMDD（4位年+2位月+2位日）- YYYYMMDD（4位年+2位月+2位日），例如：20160330-20160407。

2. 相关说明

数据内容描述说明

bibli and abstract：著录项目与文摘数据

full text：全文文本

full image：全文图像

legal status：法律状态

enriched citation：引文

family：同族

reexamination：复审

invalidation：无效

verdict：判决书

（六）数据说明

1、文件包层级结构说明

示例：YYYYMMDD-N1.ZIP/C1NNNNNNNNN/ C1NNNNNNNNN.XML

YYYYMMDD-N1.ZIP

YYYYMMDD为生产日期，如20160330； N1表示生产日期的数据按容量是否超过2G分隔的压缩包个数，每个压缩包的数据容量以2G进行划分，超过2G后顺序排列，如20160330-002.ZIP为20160330期的第二个2G压缩包…以此类推。

C1NNNNNNNNN

C1为公开国别代码（二位大写字母），NNNNNNNNN为申请号，C1NNNNNNNNN为文件夹名称。该文件夹中包括：文件名称为C1NNNNNNNNN.XML的中国引文代码化文件。

2、主要数据元素说明

元素名称	元素内容
ApplicationInfo	申请信息
PublicationInfo	公开信息
ReferencesCited	引证信息（包括丰富引文信息）
ForwardCit	被引证信息