

中国专利文献权威文档技术说明 V1.0

中国专利文献权威文档（以下简称“权威文档”）包含中国国家知识产权局 1985 年至今所公开的发明、实用新型、外观设计三种类型各个公布级的专利文献的信息。

权威文档将提供给其他知识产权机构（或其他潜在的用户），以便于评估其所获得的中国专利文献资源的完整性。权威文档的技术内容如下。

权威文档以过档数据及周更新数据两种形式进行发布：过档数据包含 1985 年 9 月 10 日至更新日的权威文档数据，每季度更新一次；周更新数据为当周公报数据的权威文档数据（与每周三公开的公报数据对应），每周/月定期更新。

1. 权威文档所包含的信息

权威文档中每件中国专利文献包含以下两方面信息。

1.1 专利文献标识

专利文献标识指在专利申请公布和专利授权公告时给予专利文献的标识信息。它包含国别代码、专利文献号、种类标识代码三部分。

（1）国别代码

根据 WIPO 标准 ST. 3，中国国家知识产权局的国别代码为 CN。

（2）专利文献号

专利文献号是指国家知识产权局按照法定程序，在专利

申请公布和专利授权公告时给予的专利文献标识号码。

权威文档中专利文献号将忠实于原始出版。

(3) 种类标识代码

种类标识代码是指国家知识产权局为标识不同种类的专利文献而规定使用的字母代码，或者字母与数字的组合代码。

权威文档中种类标识代码忠实于原始出版。

1.2 公开日期

与专利文献标识相对应的专利文献公开日期。对于申请公布类专利文献，公开日期指申请公开日、申请公布日等；对于授权公告类专利文献，公开日期指审定公告日、授权公告日等。

公开日期的格式规范为 YYYYMMDD 的形式，其中 YYYY 为 4 位数字表示的年，MM 为 2 位数字表示的月，DD 为 2 位数字表示的日。

2. 权威文档的组织方式

中国专利文献权威文档按发明、实用新型、外观设计分为三个部分，每部分对应一个文本文件，采用 UTF-8 字符编码，其中分别存储发明、实用新型、外观设计专利的信息。

2.1 文件名称

每个权威文档文件的文件名称采用“专利类型+文献范围”两部分信息，两部分信息之间用空格进行分隔，其中。

- “专利类型”取值为：Invention, utility model, design, 分别代表发明、实用新型、外观设计。

- “文献范围”为权威文档中所包含文献公开时间的起止范围，开始时间和结束时间之间用字符“-” 进行连接。

例如：“INVENTION 19850910-20121117”，表明权威文档包含了 1985 年 9 月 10 日至 2012 年 12 月 17 日所出版的中国发明专利文献的信息；“INVENTION 20160824”，表明权威文档包含了 2016 年 8 月 24 日所出版的中国发明专利文献的信息。

2.2 内部结构

每个权威文档的内部结构如下。

- 专利文献记录之间以 CRLF 进行分隔；
- 国别代码、专利文献号和种类标识代码联合使用，与公开日期之间用“;” 进行分隔；
- 专利文献记录按专利文献号排序，专利文献号相同时，按种类标识代码排序。

图 1 为一个示例。

```
CN301455224S;20110202
CN301455239S;20110202
CN301455242S;20110202
CN301455271S;20110202
CN301455291S;20110202
CN301455325S;20110202
CN301455349S;20110202
```

图 1：权威文档示例