**商标平台redis数据库设计**

商标平台针对不同的业务使用不同的redis数据库，redis数据库的redis server的IP与端口号固定，主要用于存放业务需要使用的历史商标数据，并为了辅助提高业务程序的效率而以多种存储形式进行组织。

具体设计如下：

# 近似商标检索业务

1:大类商标编号集

1）数据类型为redis的有序集合zset。大类商标编号集以目前商标局定义的四十五个大类为索引，将所有的历史商标数据按照其注册的大类编号进行聚合，每个历史商标名以“商标注册编号”拼接“商标名”的形式存入编号集合，进行去重处理，并分配一个在此大类的集合中全局唯一的递增序号。

2)注：此库中数据只为了计算近似商标服务，所存储数仅是为了服务于具体计算程序，并不作为真实历史数据的备份使用。主要目的在于利用1）中提到的递增序号来指代某一具体的商标，减少检索时做筛选过滤的时空开销。

3）key 的规则为:

brank::大类编号（不含前导0）

“商标注册编号”拼接“商标名”时，使用的分隔符为 &\*( 共三个字符。

4）以第3大类的商标编号集合的前10个数据举例，每行数据只有两个值，一个是score，这里给它赋予的含义是对应的value在集合里的编号；另一个则是value，内容为前面所述的注册编号和商标名的拼接。（简单解释的话，zset就是一个把存储的value按score排序存储和访问的集合）

brank::3 =>[

1， 17756682&\*(爱咪儿

2， 17752753&\*(JINGOVERLORD 京霸王

3， 11579014&\*(德机

4， 17729965&\*(柯美丽

5， 11563076&\*(吉婴岛

6， 17709571&\*(原筋堂

7， 17680064&\*(御春汉方

8， 11462543&\*(YESFLOWERS

9， 21968292&\*(汉古沁姿

10， 17470702&\*(德贝尔

]

2:历史商标数据库

1）历史商标数据库以具有相同前缀的关键字拼接key进行组织，数据类型为hash，含有复数个字段，每个key记录“属于某个大类的某个历史商标的数据信息”。

2)注：此库中数据只为了计算近似商标服务，字段的设置、取值均是为了服务于具体计算程序，并不作为真实历史数据的备份使用。

3）key 的生成规则为:

bData::大类编号（不含前导0）::此商标在这个大类里的序号（即此商标在大类商标编号集里对应的序号score值）

4）散列中的键和值的类型说明如下：

|  |  |  |
| --- | --- | --- |
| 键 | 类型 | 说明 |
| bid | Integer | 此商标在大类商标编号集里对应的序号score值 |
| name | String | 商标名 |
| no | String | 商标编号 |
| sts | int | 注册状态，0表示失败，1表示成功 |
| py | String | 商标的中文部分转换成的拼音字符串，以每个字的拼音为单位逗号分隔 |
| ch | String | 商标的中文部分 |
| eng | String | 商标的英文部分，以词为单位逗号分隔，并统一为小写字母 |
| num | String | 商标的数字部分，以商标名中每段连续的数字为单位逗号分隔 |
| date | String | 商标的注册时间字符串，格式为”年月日”，有前导0 |

示例：

bData::3::2 =>[

“bid”:”2,

“name”: “JINGOVERLORD 京霸王”,

“no”: “17752753”,

“sts”: “1”,

“py”: “jing1,ba4,wang2”,

“ch”: “京霸王”,

“eng”: “jingoverlord”,

“num”: “”, ##商标名中没有的部分取空串

“date”: “20150825”

]

3:历史商标商品项库

1）历史商标商品项库以具有相同前缀的关键字拼接key进行组织，数据类型为set，每个key记录“属于某个大类的某个历史商标的所有商品项编号”。

2)注：此库中数据只为了计算近似商标服务，字段的设置、取值均是为了服务于具体计算程序，并不作为真实历史数据的备份使用。

3）key 的生成规则为:

bItem::大类编号（不含前导0）::此商标在这个大类里的序号（即此商标在大类商标编号集里对应的序号score值）

4） 以第1大类的商标编号集合的序号为1的数据举例，key对应的数据即为1大类中序号为1的商标，在1大类里都注册了哪些商品项的序号（并不一定是纯数字串）。

bItem::1::1 =>[

010030

010573

010002

010132

010666

010232

010078

C010251

010274

010584]

4:大类商标名读音集

1）大类商标名读音集以具有相同前缀的关键字拼接key进行组织，数据类型为set，每个key记录“属于某个大类中含有某些汉字读音/英文单词的商标的大类内序号”。

2)注：此库中数据只为了计算近似商标服务，字段的设置、取值均是为了服务于具体计算程序，并不作为真实历史数据的备份使用。

3）key 的生成规则为:

bPySet::大类编号（不含前导0）::若干个汉字读音/英文单词的逗号拼接串

4） 在前文介绍的历史商标数据库的例子的基础上举例，以第3大类的，同时含有拼音为jing1（比如汉字“京”）、拼音为wang2（比如汉字“王”）的商标的编号集合为

bPySet::3::jing1,wang2 =>[

2

3422]

为了验证，相应的去查看 bData::3::2和bData::3::3422，有：

bData::3::2 =>[

“bid”:”2,

“name”: “JINGOVERLORD 京霸王”,

“no”: “17752753”,

“sts”: “1”,

“py”: “jing1,ba4,wang2”,

“ch”: “京霸王”,

“eng”: “jingoverlord”,

“num”: “”, ##商标名中没有的部分取空串

“date”: “20150825”

]

和

bData::3::3422 =>[

“bid”:”3422,

“name”: “白鲸王”,

“no”: “21862385”,

“sts”: “1”,

“py”: “bai2,jing1,wang2”,

“ch”: “白鲸王”,

“eng”: “”,

“num”: “”,

“date”: “20161110”

]

两条数据，即bPySet::3::jing1,wang2间接指向了一个集合[JINGOVERLORD 京霸王， 白鲸王]，集合里的商标都有同音字jing1和wang2。