名词解释：

A类最佳匹配位：彩铃专区（彩铃、视频彩铃、音频彩铃）

B类最佳匹配位：歌手、专辑、电台、期刊、视频。

**V2.7.0**

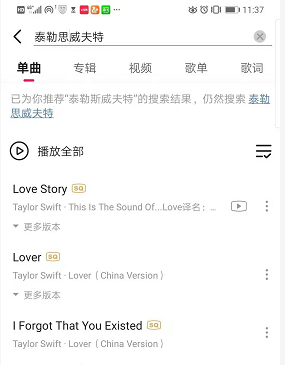
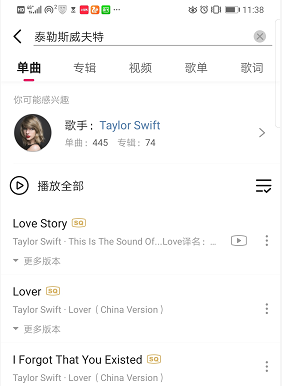
1. **最佳展示位支持匹配纠错词**

（纠错词也支持匹配最佳展示位，纠错词和原词的逻辑保持一致。注意：纠错结果类型有：歌手、专辑、歌曲、tag、影视、综艺 6类型，但只有歌手、专辑有最佳展示，其他类型的最佳展示不受纠错词影响）

【需求描述】

原有：当搜索词有纠错结果时，不支持匹配最佳展示位，如搜索“泰勒思威夫特”，最佳展示位无“泰勒斯威夫特”歌手；

改为：纠错词也支持匹配最佳展示位，纠错词和原词的逻辑保持一致。如纠错结果为“beyond”，最佳展示位仍然匹配歌手beyond，而不是专辑beyond。

**2.普通歌曲组增加影视字段匹配（歌曲组搜索时扩展影视名字段的匹配；影视组搜索时，增加过滤条件 movieHotLevel=6）**

【需求背景】

音乐库中存在不少与影视名称同名的非热门歌曲，由于与搜索词精确匹配导致排序靠前，如“长安十二时辰”，同名非热门的歌曲排在第一位。

【需求描述】

歌曲搜索时影视名优先级等同于歌曲名优先级，即匹配歌名“长安十二时辰”和匹配影视名“长安十二时辰”的得分是一样的，根据热度进行排序。

另，如果匹配的影视名是热门影视（影视热度是6），优先级仍然最高。

**3.优化歌词和歌曲同名时的搜索效果**

【需求描述】

当前：当搜索结果中有歌名全匹配结果时，不再匹配歌词。

改为：若全匹配的歌曲热度小于10万，仍然搜索歌词，且歌曲和歌词无优先级，按热度自动排序。



**4.演唱会AI剪辑和人工剪辑内容去重**

~~【需求背景】~~

~~对于演唱会和大型音乐活动，会用AI剪辑出一部分现场短视频，后续会有人工剪辑，质量优于现场短视频。~~

~~人工剪辑和AI剪辑内容都属于MV素材中的现场短视频分类。~~

~~【需求描述】~~

1. ~~对于AI剪辑内容，给定新视频标识，标识3小时后失效；~~
2. ~~对于人工拆条内容，给定新视频标识，标识48小时后失效；~~
3. ~~优先级：人工拆条（有新视频标识） > AI剪辑内容（有新视频标识） > MV > 普通现场短视频 > PGC视频 > UGC视频。~~

~~注： AI剪辑和人工拆条内容通过CPID区分,AI剪辑CPID（699223），人工拆条（699119）~~

改为：

屏蔽MV素材中现场短视频分类里AI剪辑内容（CPID=699223），屏蔽指的是下架，可人工恢复上架。（屏蔽MV素材中现场短视频分类里AI剪辑内容（CPID=699223），屏蔽指的是下架，可人工恢复上架

注意：可以做成配置项，支持多值）

**5.支持咪咕音箱的语义搜索**

【需求描述】

当搜索词包含汉字，且搜索词长度大于等于5时，可以判断是普通搜索词还是语义搜索词，可以从语义搜索词中提取关键词以及关键词类别。

1. 搜索词包括单个关键词

* 单个关键词：歌曲名、歌手名、词曲作者、专辑名、影视/综艺名、标签名、歌词

1. 组合关键词

* 组合关键词1：歌手名+歌曲名、歌手名+专辑名、歌手名+综艺名、歌手名+标签名、标签名+标签名；
* 组合关键词2：单个关键词+排序、组合关键词+排序。

**【举例】**

1. 单个关键词：

* 输入“我想听花香”，输出关键词：花香，关键词类别：歌曲；
* 输入“推荐歌手李白的歌曲”，输出关键词：李白，关键词类别：歌手；
* 输入“我想听钢琴曲”，输出关键词：钢琴，关键词类别：标签。

1. 组合关键词1：

* 输入“我想听周杰伦的告白气球”，输出关键词：周杰伦，类别：歌手；关键词：告白气球，类别：歌曲；
* 输入“播放陈奕迅的粤语歌”，输出关键词：陈奕迅，类别：歌手；关键词：粤语，类别：标签。

1. 组合关键词2：

* 输入“周杰伦的最新单曲”，输出关键词：周杰伦，类别：歌手；排序规则为按时间倒排。
* 输入“播放陈奕迅的最热粤语歌曲”，输出关键词：陈奕迅，类别：歌手；关键词：粤语，类别：标签；排序规则为按热度倒排。

**6.支持歌词置顶功能**

【需求描述】

歌词搜索接口增加搜索置顶功能。当进行歌词搜索时，若用户搜索词与运营平台配置的置顶搜索词精确匹配时，优先返回配置内容。

**7.A类最佳匹配位场景扩展【涉及前端联调】**

* 1. **最佳匹配位增加【票务】类型**

【需求描述】

1. 当前无票务数据同步，所以仅需要支持人工配置即可；
2. 运营管理平台增加【票务】的最佳展示位干预功能，搜索接口返回最佳展示位类型和配置的票务id即可；
3. 优先级：票务>彩铃专区。如果运营平台既配置了彩铃，又配置了票务，搜索接口优先返回票务内容。
   1. **最佳匹配位【演唱会】类型取消限制**

【需求描述】

原有：只匹配直播状态演唱会；

改为：

1、对于人工干预的，可匹配直播中、预约、完整回放三种状态的演唱会；对于自动匹配的，仍然限制直播状态（需要门户再进行一次状态校验）~~。另，最佳展示位取消通过匹配 地名/歌手 召回演唱会。优~~先级：演唱会>票务>彩铃专区。

**8.视频搜索支持拆条类型【涉及前端联调】**

【需求描述】

从cms的短视频反向查询接口获取人工切条数据，方式：组合查询“短视频来源”=统一CMS,和“播放源类型”=切片类型。

在视频tab搜索和歌手下MV搜索时，支持匹配该数据源数据，且优先级高于其他MV素材中的各类数据（MV、现场短视频、PGC等）。

**9.支持视频彩铃的渠道过滤**

【需求描述】

运营平台增加视频彩铃的渠道过滤，对于已被过滤的渠道，不返回对应的视频彩铃内容。

1. **组合无关联搜索优化**

【需求背景】

当前，搜索词为歌手+歌曲/专辑组合无关联时，如“周杰七里香”，会判断歌手热度是否大于10万，若不满足，则会出现搜索无结果。

【需求描述】

对上述歌手+歌曲/专辑组合无关联场景，同时判断歌曲/专辑热度，若歌曲/专辑热度大于10万，也需要根据歌曲/专辑名进行搜索。

**10.运营平台**

* 1. **支持歌词搜索置顶配置**

【需求描述】

1. 推荐管理——>置顶配置，增加歌词类型。



* 1. **渠道过滤支持配置视频彩铃**

【需求描述】

内容管理——>渠道过滤。增加视频彩铃类型，页面展示主键、名称、歌手、渠道、更新时间、操作控件。



* 1. **支持内容提前屏蔽**

【需求背景】

重点保障时，专辑和歌曲内容不允许提前露出，当内容未分发时，目前只能通过设置屏蔽词来保障，会对一些正常内容有影响。

【需求描述】

1. 渠道过滤支持内容提前屏蔽，即当内容未入库时可提前在运营管理平台配置渠道过滤；
2. ~~查看页面增加【高危】筛选项，若选中该选项，点击搜索时，可将所有已配置，且在内容表中未查询到的内容筛选出来。~~



* 1. **最佳展示位干预增加“票务”类型**

【需求描述】

推荐管理——>最佳展示位干预，增加“票务”类型，由于没有真实数据同步，所以只需要将运营配置的票务id返回给搜索平台即可，不需要去数据库校验是否存在该id。



**11.纠错**

* 1. **纠错分级**

【需求背景】

当前是优先匹配度>热度进行纠错结果的选择。音乐由于内容比较多，存在相似名称的场景比较多。在匹配度相差不大的情况下，希望能返回热度更高的内容作为纠错结果。

【需求描述】

增加纠错范围得到召回结果取热度值最高的一个，与当前纠错结果进行对比，若前者热度比后者大10万及以上，或者hotlevel前者比后者大1，则选择前者作为最终的纠错结果，否则，选择后者作为最终的纠错结果。

badcase：如下图，搜索“刚刚好遇到你”，盈妹妹的“刚刚好遇到”热度在10万左右，李玉刚的“刚好遇见你”热度在45万左右，两者和搜索词的编辑距离相差只有1，热度值差了4倍多，此时，希望优先返回热度高的这个作为纠错结果。



* 1. **影视名称支持入纠错词典**

【需求描述】

对于热度大于0的歌曲，同时将其对应的影视名称加入纠错词典，热度为歌曲热度减1000（最小为0）。当歌曲下架之后，对应的影视名称也下架。

注意：一个影视名称可能对应多个歌曲，需要考虑如何处理影视名称在纠错词典中的上下架问题。

* 1. **热门歌手/歌曲的别名也入纠错词典**

【需求描述】

Hot> 10万或者hotlevel > 1的歌曲和歌手，别名支持加入纠错词典。对于个数少于5的，只支持同音纠错，大于等于5的还支持错/少/多一个字纠错。

badcase：“玛丽莲慢森”不支持纠错为“玛丽莲曼森”。

* 1. **白名单不纠错，汉字入纠错词典可准实时更新**

【需求描述】

1. 运营管理平台增加白名单配置功能，对于白名单中的词，不进行纠错处理。
2. 运营管理平台增加纠错词典配置功能，可准实时更新。

**12.爬虫**

【需求描述】

1. 对于外文歌手和歌曲，爬虫支持匹配译名、别名

由于部分外文歌手和歌曲与竞品存在原名不同，导致热度爬取时无法匹配。改为可用我们数据中的别名/译名数据同时与竞品匹配。

1. 增加热门歌手的歌曲热度爬取

歌手热度限制大于5万。

# V2.7.2

名词解释：

CPID：版权id前6位为cpid。

1. **数据处理支持对指定的CPID进行屏蔽和放开**

【需求背景】

音乐计划从第三方CP引入600万条数据，分批同步到搜索，在12月31号之前，这批数据不支持搜索。这批数据有固定的几个cpid。

【需求描述】

1、数据处理增加cpid配置，对于满足配置的CPID的数据，入库时强制设置为下架状态；

2、开关开启之后，这批数据陆续设置为上架状态。

1. **最佳匹配位票务数据增加来源**

【需求描述】

1. 运营平台。最佳匹配位支持增加票务来源配置，通过筛选进行选择，筛选项包括：己方项目（00）、永乐项目（01）。批量操作模板中也需要可编辑来源。
2. 搜索接口。咪咕音乐的票务数据目前来源于自有、第三方两个来源，最佳匹配位返回票务数据时需要指明当前票务的来源。
3. **数据处理日志优化**

【需求描述】

根据“统一搜索项目日志规范”输出可供ES采集的标准接口日志信息。

日志示例：

{"moduleName":"elt-music","classMethod":"syncWirelessProduct.json","requestParam":"{\"material\":\"{\\\"songName\\\":\\\"爱的可能\\\",\\\"nameAlias\\\":\\\"\\\",\\\"versionName\\\":\\\"60000292018041315400901184155\\\",\\\"versionInfo\\\":\\\"60000292018041315400901184155\\\",\\\"songNamePinyin\\\":\\\"aidekeneng\\\",\\\"length\\\":\\\"00:04:26\\\",\\\"translateName\\\":\\\"\\\",\\\"artistList\\\":[{\\\"name\\\":\\\"叶倩文\\\",\\\"artistId\\\":1000000499}],\\\"tags\\\":[{\\\"id\\\":1000001672,\\\"tagName\\\":\\\"流行\\\"}],\\\"intro\\\":\\\"\\\",\\\"albumList\\\":[{\\\"albumName\\\":\\\"出口\\\",\\\"trackNumber\\\":4,\\\"albumId\\\":15554,\\\"disc\\\":\\\"1\\\"}],\\\"songNameFirstLetter\\\":\\\"ADKNADKNAD\\\",\\\"composerList\\\":[{\\\"name\\\":\\\"李偲菘\\\",\\\"artistId\\\":1000000902}],\\\"songId\\\":6,\\\"releaseDate\\\":\\\"2019-02-28\\\",\\\"lyricistList\\\":[{\\\"name\\\":\\\"李偲菘\\\",\\\"artistId\\\":1000000902}],\\\"lyricUrl\\\":\\\"http://files/lyric/2017-01-11/1701111340588651.lrc\\\"}\",\"action\":\"2\",\"PID\":\"6000029\",\"transactionId\":\"6\",\"SEQ\":\"60000292018041315400901184155\",\"KEY\":\"94aaddcba6514d71e6dd99035fc5139e\",\"objectType\":\"1\"}","timeCost":"1000","resultNum":"1","result":[{"resultCode":"0000000","programId":"656637494","resultMessage":"success"}]}

1. **问题处理**

* 问题现象：玛丽莲慢森，无法纠错为“玛丽莲曼森”，问题原因：cn\_name未入纠错词库。测试时注意人物和歌曲的cn\_name 都检查下。

# V2.8.0

名词解释：

S类最佳匹配位：活动/专区

A类最佳匹配位：彩铃专区（彩铃、视频彩铃、音频彩铃）

B类最佳匹配位：歌手、专辑、电台、期刊、视频。

**1.独立票务搜索**

【需求描述】

咪咕购票平台页面，增加独立的票务搜索。增加独立票务搜索接口，支持搜索出咪咕的票务数据。

**【参与搜索字段】**

场次名称、演出城市、演出地点、演出分类、艺人名称；

**【参与筛选的字段】**

演出城市、票务类型（选座、不选座）、演出分类；

**【排序规则】**

1. 支持人工置顶。当搜索词与运营配置搜索词精确匹配时，优先返回人工配置的内容；
2. 精确匹配 > 前中缀匹配 > 模糊匹配；
3. 演唱会>音乐会>话剧舞台剧>亲子；
4. 前两者相同情况下按艺人热度排序；相同艺人且同匹配度情况下，按演出时间排序，优先展示最近场次；

**【其他规则】**

1、搜索平台需返回票务场次id和来源（00：自有；01：永乐）；

2、已下架项目不支持搜索；

3、场次名称、演出城市地点、艺人名称不重复加分。

1. **歌单内歌曲搜索、歌手主页内搜索**

【需求描述】

1. 支持指定歌单id和歌单内歌曲名称，进新歌单内歌曲搜索，与歌曲搜索接口匹配和排序逻辑保持一致即可。
2. 支持指定歌手id和名称进行歌手主页内歌曲、专辑、MV搜索，与SAM接口匹配和排序逻辑保持一致即可。
3. **语义搜索**

【需求描述】

当搜索词包含汉字，且搜索词长度大于等于5时，可以判断是普通搜索词还是语义搜索词，可以从语义搜索词中提取关键词以及关键词类别。

1. 搜索词包括单个关键词

* 单个关键词：歌曲名、歌手名、词曲作者、专辑名、影视/综艺名、标签名~~、歌词~~；

1. 组合关键词

* 组合关键词1：歌手名+歌曲名、歌手名+专辑名、歌手名+综艺名、歌手名+标签名、标签名+标签名；
* 组合关键词2：单个关键词+排序、组合关键词+排序。

1. 支持语义同义词

支持在运营管理平台配置语义同义词，搜索接口支持去运营配置，如：

* 给标签“钢琴”配置同义词“钢琴曲”，可通过“钢琴曲”搜索出带有“钢琴”标签的歌曲；
* 给“舒缓”配置“驰放|抒情流行”的同义词，可搜索出带有“驰放”或者“抒情流行”标签的歌曲。

\*注：

最佳展示位、无维度、歌曲、专辑、视频、~~歌词~~均需要支持语义搜索。

**【举例】**

1. 单个关键词：

* 输入“我想听花香”，输出关键词：花香，关键词类别：歌曲；
* 输入“推荐歌手李白的歌曲”，输出关键词：李白，关键词类别：歌手；
* 输入“我想听钢琴曲”，输出关键词：钢琴，关键词类别：标签。

1. 组合关键词1：

* 输入“我想听周杰伦的告白气球”，输出关键词：周杰伦，类别：歌手；关键词：告白气球，类别：歌曲；
* 输入“播放陈奕迅的粤语歌”，输出关键词：陈奕迅，类别：歌手；关键词：粤语，类别：标签。

1. 组合关键词2：

* 输入“周杰伦的最新单曲”，输出关键词：周杰伦，类别：歌手；排序规则为按时间倒排。
* 输入“播放陈奕迅的最热粤语歌曲”，输出关键词：陈奕迅，类别：歌手；关键词：粤语，类别：标签；排序规则为按热度倒排。

1. **最佳匹配位优化**
   1. 支持多值（涉及客户端联调）

【需求描述】

运营管理平台进行最佳展示位配置时支持id多值，搜索接口对于同一类型的配置，支持返回多值，如：

* 在给“5G”配置视频彩铃最佳展示位时配置多个视频彩铃id，则搜索接口支持将所配置的id均返回前端；
  1. 最佳匹配位S区增加活动/专区（涉及客户端联调）

【需求描述】

运营管理平台支持活动/专区配置，搜索接口与配置关键词精确匹配时，返回配置的url地址，如：

* 给“咪咕汇”配置购票活动url。
  1. 支持热门歌手/专辑前中缀匹配

【需求描述】

对于热门歌手/专辑（hot > 20万 或者 hotlevel >0）, 支持前中缀匹配：当同时存在前中缀匹配结果和单一纠错结果时，对比两者搜索结果的热度，优先选择热度高的；如搜索“林俊”，前缀匹配出“林俊杰”、纠错匹配出“林军”，则优先选择结果“林俊杰”；

* 1. ~~票务数据支持返回票务来源~~

~~【需求描述】~~

~~咪咕音乐的票务数据目前来源于自有、第三方两个来源，最佳匹配位返回票务数据时需要指明当前票务的来源。~~

1. **视频彩铃搜索优化**

【需求描述】

1. 支持歌手名、歌手+彩铃名称组合搜索：

* “周杰伦的视频彩铃”；
* “周杰伦的视频彩铃说好不哭”；
* “视频彩铃小精灵TFBOYS”。

1. **专辑搜索优化**

【需求描述】

* 1. 基本排序规则变更

当前排序：专辑名称 > 歌曲名称， 时间 > 热度。

改为：热门专辑组（hot >10万或者hotlevel>0，要求有版权歌曲数量大于0） > 其他组。各组内按时间倒排，时间相同按热度倒排。

badcase：搜索“芒种”，音阙诗听的芒种排序靠后。“刚好遇见你”，歌曲全部无版权的罗艺恒的专辑排在前面。

* 1. 纠错逻辑优化

根据客户端调用的接口，在调用纠错接口时指定当前搜索类型（如专辑、MV），纠错系统根据该类型调整纠错结果（参见17.4纠错优化）。

* 1. 组合模糊搜索优化

当前英文组合模糊搜索时，由于数字专辑的加分规则，容易导致匹配度较低的数字专辑排序靠前。需要修改为，模糊搜索时，优先按照匹配度得分，然后再根据数字专辑优先于普通专辑的的逻辑进行处理。

badcase: “Be There for You  **Meghan Trainor**”、“Nice To Meet **Niall Horan**”

1. **影视名称支持配置别名**

【需求描述】

运营管理平台支持影视别名配置，影视别名与原名匹配优先级相同，且共用movielevel的等级。

badcase：“哈利波特”搜不出影视“哈利波特与阿兹卡班”歌曲。“三生三世”，影视“三生三世十里桃花”的配乐排序靠后。

1. **组合精确搜索时，如果结果无版权，支持单一实体匹配**

【需求描述】

当前：当组合精确搜索有结果时，只返回该精确结果。

改为：当组合精确搜索有结果时，若该精确结果无版权，则支持使用单一实体进行匹配（只召回热度大于10万，且有版权的歌曲），召回结果排在组合精确结果下方。

badcase：“我们的歌 王力宏”。

1. **MV纠错逻辑调整**

【需求描述】

当前：纠错词不进行前中缀匹配，导致很多MV搜索时纠错后无结果。

改为：当纠错词类型为歌曲、专辑时，支持用该纠错词匹配MV名称，支持前中缀匹配，当纠错词类型为歌手时，支持用该纠错词匹配MV的歌手信息，支持精确匹配。

badcase：aoxiang天地、“Truth Herts”、“不成时”、“wohowodezuguo“

1. **联想提示优化**

【需求描述】

1. 联想无结果时支持数字汉字转换

badcase：7里香。

1. 数据处理TOP10000入联想词典时，需要检查这些词是否是热门歌手名称，是的话，给与热门歌手的标识。搜索接口从联想词库里取提示词时，对于有热门歌手标识的，返回对应模块，支持前端搜索直达（如搜索周杰伦时，不用点击搜索，直接跳转结果页）。
2. **抽屉排序优化**

【需求描述】

原：isCopyright > versionType > hot；

改为：isCopyright > versionType > sq音质 > hotlevel\*100,000+hot > 存在有关联有版权MV > 其他；

优化原因：主要是为了在歌曲无版权的时候可以增加有版权MV的露出。

先录音棚、现场版

sq音质的用起风了来看

2个歌曲，一个

1. **www打乱式排序优化**

【需求描述】

无维度搜索当前支持“抽屉式”和“打乱式”两种排序，若请求参数中指定“打乱式”排序，默认只保留抽屉最外围内容，抽屉里面的内容都自动降权排在后面。（精确 > 分词模糊，有版权 > 无版权）

1. **~~支持对指定的CPID提供搜索开关~~**

~~方案待定。~~

1. **组合搜索高亮优化**

【需求描述】

对于纠错接口返回的组合有关联和组合无关联的结果，需要将组合中各个实体的名称都作为高亮词返回给前端。

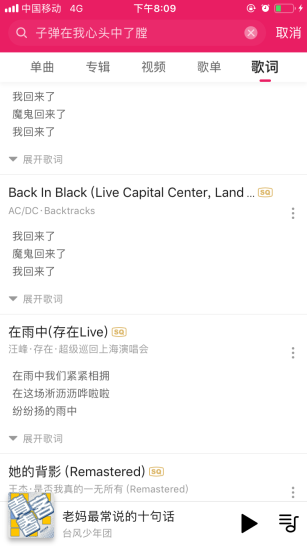
Badcase：“周杰伦 青花瓷”，返回的高亮词只有一个原词“周杰伦青花瓷”，导致前端“周杰伦”和“青花瓷”都未高亮。

1. **歌词模糊搜索优化 – 2.6.0遗留**

【需求描述】

当前：根据分词匹配度进行歌词搜索结果排序。导致搜索“子弹在我的心头中了膛”，歌词匹配度更高的歌曲“你的酒馆对我打了烊”排序靠后。

改为：可以综合考虑匹配度和编辑距离。当前排在前面的可能匹配到的分词个数更多，但是较分散，编辑距离比较大，不符合用户搜索行为习惯（搜索歌词时更多的是输入一句连续的歌词），排序应该靠后。



1. **音箱搜索优化**

【需求描述】

1. 语义搜索支持标签同义词，参见需求【3：语义 搜索】中的标签同义词需求描述；
2. 影视名称支持匹配多值，与app共用同一个配置。
3. **纠错优化**
   1. 支持汉字+简拼转简拼纠错

【需求描述】

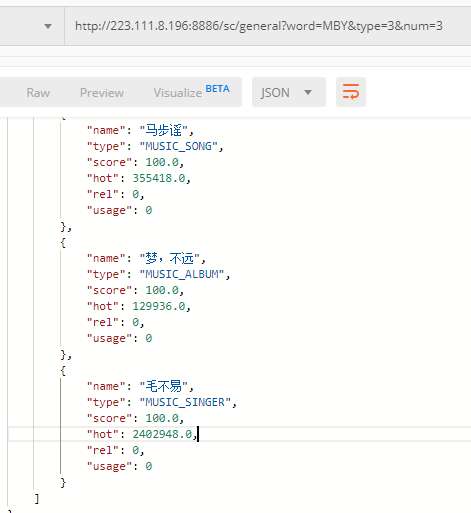
当前：当搜索词包含汉字时不支持转简拼纠错，如“cx坤”，不支持纠错为“蔡徐坤”、“歌唱zg”，不支持纠错为“歌唱祖国”。

改为：如果搜索词包含汉字+字母，支持转简拼进行单一纠错。当转简拼单一匹配上（纠错类型为0），组合有关联也匹配（纠错类型为1），选择两类中最高热度最大的结果，热度相同时，优先取组合有关联作为纠错结果。

* 1. 简拼纠错优化

【需求背景】

对于简拼来说，纠错接口经常会返回多个不同的纠错词，如“MBY”。搜索接口会将前三个纠错结果逐一降权匹配，容易导致纠错提示和实际搜索不匹配。

【需求描述】

对于简拼搜索（组合搜索不支持简拼），当返回多个不同的纠错词时，只取热度最大的纠错词对应的结果，如果该纠错词对应多个类型，则各个类型都返回，按热度降序。如：MBY还有一个纠错结果是“毛不易”影视名，则搜索“MBY”时，纠错接口返回“歌手毛不易”和“影视名毛不易”两个，歌曲“马步谣”、专辑“梦，不远”则过滤掉。

* 1. 优化组合搜索存在单字时的纠错效果

【需求描述】

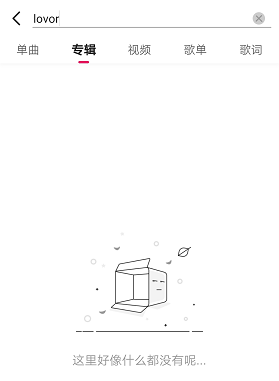
当前：很多单字不入纠错词典，导致单字组合搜索时经常会出现错误纠错结果。要求单字也正常入纠错词典，若组合纠错结果与其他类型纠错词相同，则优先取其他类型纠错词。如搜索“爱陈慧琳”，满足组合有关联“歌手陈慧琳 歌曲爱”，若同时有个歌曲就是“爱陈慧琳”，则优先取“歌曲 爱陈慧琳”的结果。

badcase:“爱 陈慧琳”、“飞 柯有伦”等

* 1. 支持指定类型纠错

【需求描述】

当搜索接口调用纠错接口时，若指定了搜索类型，则当存在非该类型的白名单时，需要继续检查是否存在该类型的纠错结果，如指定专辑，搜索“lovor”，虽然对于歌曲，“lovor”是白名单，但是对于专辑并不是，所以需要获取专辑下的纠错结果“lover”。对比“歌曲lovor”和专辑“lover”的热度，取热度最高的一个返回给搜索系统。



* 1. 歌手入纠错词典增加限制

【需求描述】

当前：歌手入纠错词典限制规则是：歌手有热度或者歌手有作品。

改为：歌手有热度且有版权歌曲数>0 或者hotlevel>0。

1. **数据处理**
   1. 票务数据同步

【需求描述】

* 自有数据：搜索平台、思特奇双方约定先同步全量自有数据。后续思特奇针对状态变更的数据（上级、下架、修改）需要进行实时通知；
* 三方数据：每日定时同步全量数据。针对平台自行发起手动变更的三方数据需要实时更新到搜索平台。
  1. 产品判断改为用STATUS【改为12月底】

【需求描述】

原有：根据中音分发的产品操作类型action字段来判断产品状态；

改为：曲库各产品接口增加产品状态字段status，搜索此字段来识别产品状态。

* 1. 同步歌单的播放量数据
  2. MV播放量统计问题处理

1. **运营管理平台**
   1. 票务最佳展示位干预支持配置票务来源—272实现

【需求描述】

最佳匹配位支持增加票务来源配置，通过筛选进行选择，筛选项包括：己方项目（00）、永乐项目（01）。批量操作模板中也需要可编辑来源。

* 1. 最佳匹配位增加活动/专区配置

【需求描述】

推荐管理——>最佳展示位干预菜单

增加活动/专区选项，该选项与视频彩铃同级。对于活动/专区，搜索词、图片url、跳转url、名称四项是必填的。



* 1. 支持语义词典管理-关联需求3 - 改为290

【需求描述】

支持语义词典修改（参考咪咕视频）

* 可新增词典，需指定该词对应的属性（歌曲、歌手、专辑等等）；
* 可删除或修改词典中已有内容（名称、属性、同义词均支持修改）。
  1. 支持热门搜索配置（参考视频）

【需求描述】

1. 默认搜索词各分类栏目可自定义扩展选项（当前仅有热门搜索），实时生效；
2. 每个栏目下的热门搜索内容支持独立配置热搜词，热搜词支持配置关联推荐词、跳转url、角标、热搜趋势。关联推荐词用于展示在app对应热搜词下方用于进行热搜词推荐，跳转url用于在用户点击热搜词后直接跳转h5页面。（推荐词、跳转URL、角标、热搜趋势都不是必填的）。
3. 需要权限较高的人点击确认后，分类热词配置才能生效。
   1. 影视名称支持多值

【需求描述】

1. 歌曲别名设置时，若新增一项或者修改已有的，下方截图中的影视名输入框需展示已有的影视名称，支持多值，运营可在原有的基础上就行修改，点击保存后覆盖原有影视名。
2. 问题处理：当前不填入影视名整个就无法保存，影视名不是必填字段。



* 1. 内容下架支持输入下架原因

【需求描述】

内容管理——>上下架菜单，点击下架或者通过模板批量下架时，支持输入下架原因，默认为操作人。

* 1. 歌单支持查看歌单下歌曲id

【需求描述】

数据查看——>歌单。

双击id进入详情页可以查看歌单内歌曲id列表。

1. **其他**

【需求描述】

~~修改咪咕音乐APP2.0 搜索接口加密用的key值。~~

* 1. 音乐门户热词和搜索接口增加熔断机制

【需求描述】

门户搜索接口调用音乐详情接口时，超时熔断时直接返回搜索数据。且在音乐详情接口完全不可用时，支持手动降级处理，直接返回搜索数据。超时时间

热词接口调用音乐热词接口时，超时熔断时返回缓存结果。且在服务启动时能够加载配置文件中热搜词。

* 1. 音乐纠错服务增加降级开关

【需求描述】

音乐纠错服务器支持如下两个服务降级配置：

1. 旧服务开关：关闭后，所有包含字母的输入，直接返回空。配置后将影响所有包含字母的纠错。

2. 新服务开关：关闭后，若新流程没有返回结果，不走旧流程纠错，直接返回空。配置后将影响汉字纠错和汉字+数字纠错。

* 1. 音乐搜索服务增加降级措施和开关

【需求描述】

1. 给搜索服务增加纠错开关，在过载时通过http api对部分搜索节点关闭纠错功能；
2. 给搜索服务增加服务降级功能及开关，在过载时通过http api对部分节点的歌曲搜索用单个查询语句完成歌曲、歌手、专辑、歌曲歌手、影视的精确及前中缀查询，不支持分词查询；
3. 线程池及httpclient连接池独立配置优化；
4. 调整搜索本地缓存数量从1万到1.5万；
5. 多solr集群方案验证。

【后续版本】

增加redis缓存、调整缓存时长（配合缓存失效检测策略）。

1. **问题处理**
2. 搜索“云朵 云朵”，歌手“云朵”的歌曲“云朵”，排序靠后，此类歌手歌曲同名的组合搜索效果都不好；
3. 搜索 “蔡徐坤 1”，蔡徐坤的专辑“1”，排序靠后；
4. 搜索不稳定，重复搜索“Sobala”，结果会变来变去。

# V2.9.0

名词解释：

A类最佳匹配位：彩铃专区（彩铃、视频彩铃、音频彩铃）、演唱会、票务、活动/专题。

B类最佳匹配位：歌手、专辑、电台、期刊、视频、歌单。

**1.歌单搜索支持返回命中歌曲名称 --门户、搜索平台、前端**

【需求描述】

歌单搜索接口中，若是通过歌曲名称召回的歌单，增加返回歌曲名称，用于前端在页面上展示。

**2.最佳展示位支持返回歌单数据 -- 运营平台、门户、搜索平台、前端**

【需求描述】

运营管理平台最佳展示位干预增加歌单类型，搜索平台支持返回该类型，用于前端在最佳展示位B类展示。

搜索平台：返回歌单id即可；

搜索门户：根据id返回歌单详情，包括：歌单名称、封面url、歌曲数、播放量（XXX.X万，保留1位小数即可）。

优先级与期刊相同，条件完全相同时优先返回歌单。



**3.最佳展示位支持返活动/专题的副标题 -- 运营平台、门户、搜索平台、前端**

【需求描述】

运营管理平台最佳匹配位干预增加活动/专题的副标题，搜索平台返回给前端即可。

**4.最佳展示位优先级规则调整 -- 搜索平台**

【需求描述】

对于B类最佳展示位，原有优先级为：普通专辑=视频。

改为：若 视频热度+100,000 > 专辑热度，则优先视频，否则优先返回专辑。

【热度计算】

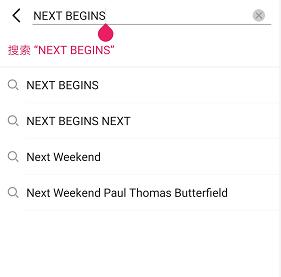
hotLevel \*100000 +hot

原有： 数字专辑 > 歌手 > 普通专辑 = 视频 > 电台 > 期刊 = 歌单。

**5.联想提示和搜索高亮忽略空格**

【需求背景】

当前，前端不支持与搜索返回的高亮词去空格匹配进行高亮，举例：1）联想返回高亮词不包含空格，包含空格的联想提示未高亮；2）搜索返回的高亮词带空格，不含包空格的搜索结果内容未高亮。如下图：

【需求描述】

搜索系统返回的高亮词和待匹配的结果进行高亮判断时，忽略两者中的空格：当高亮词不带空格，待匹配结果包含空格时，只要两者字母/汉字能够全匹配，则展示高亮；同理，当高亮词带空格，待匹配结果不包含空格时，也需要能够去空格匹配高亮。

**6.过滤部分演唱会和票务的状态**

【需求描述】

对于票务：若状态不属于“售票中、预约中、预售中、已售罄”这其中的任何一种状态，则不返回票务状态，前端不展示状态信息；

对于演唱会：若状态不属于“直播中、预约、完整回放”这其中的任何一种状态，则不返回演唱会状态，前端不展示状态信息。

**7.问题处理**

【需求描述】

MV存量热度更新。

# V2.10.0

**1.热门歌曲、专辑、歌曲支持中缀匹配**

【需求描述】

1、对于热度大于10万或者hotlevel>0的歌曲、专辑、MV，支持中缀匹配，中缀匹配优先级等同于分词全匹配。

badcase：“环环”搜不到“世界与你环环相扣”。（测试时注意取消置顶配置） ；

**2.纠错词和原词相同时，最佳展示位仍然要支持前缀匹配**

【需求描述】

当前： 当搜索词和纠错结果相同时，最佳展示位只做了精确匹配，未支持前缀匹配，导致搜索“taylor”,最佳展示位未匹配到Taylor swift；搜索“周杰”，最佳展示位未匹配周杰伦；搜“微微一笑很倾城”，热度更高的专辑“《微微一笑很倾城》电视原声”未在最佳展示位等问题。

改为：不管纠错结果和搜索词是否相同，最佳展示位都支持前缀匹配，然后按照热度来判断是选择前缀匹配结果还是选择纠错词匹配的结果。

**3.抽屉排序支持至臻音质优先**

【需求背景】

咪咕音乐新增了更高音质的24bit至臻音质。后续可能会有更高音质。

【需求描述】

原 有：isCopyright > versionType > sq音质 > hotlevel\*100,000+hot > 存在有关联有版权MV > 其他；

改为：isCopyright > versionType > 音质（24bit至臻音质> sq音质）> hotlevel\*100,000+hot > 存在有关联有版权MV > 其他；

**4.最佳展示位组合无关联增加限制**

【需求描述】

原有：对于纠错类型3（拆词） 时，最佳展示位不使用纠错结果进行匹配。

改为：原有逻辑保持，另增加约束如下：当纠错类型为2 时，若 纠错结果不是 歌手+专辑+X，歌手+歌曲+X，则不使用该纠错结果匹配最佳展示位。

**5.去除联想直达功能**

【需求描述】

在联想提示中去除热门歌手标识，前端取消歌手联想直达功能。

**6. 2.8.0版本遗留的若干纠错功能**

**6.1支持汉字+简拼转简拼纠错**

【需求描述】

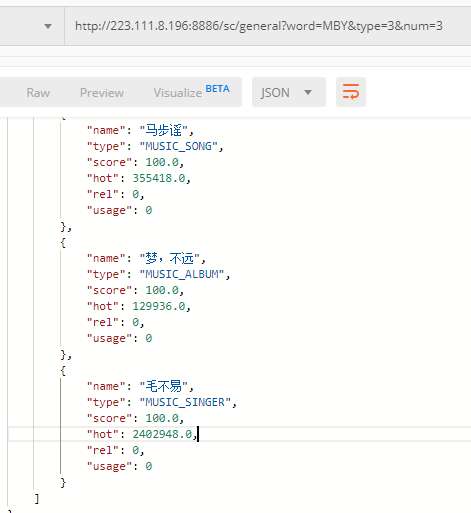
当前：当搜索词包含汉字时不支持转简拼纠错，如“cx坤”，不支持纠错为“蔡徐坤”、“歌唱zg”，不支持纠错为“歌唱祖国”。

改为：如果搜索词包含汉字+字母，支持转简拼进行单一纠错。当转简拼单一匹配上（纠错类型为0），组合有关联也匹配（纠错类型为1），选择两类中最高热度最大的结果，热度相同时，优先取组合有关联作为纠错结果。

**6.2简拼纠错优化**

【需求背景】

对于简拼来说，纠错接口经常会返回多个不同的纠错词，如“MBY”。搜索接口会将前三个纠错结果逐一降权匹配，容易导致纠错提示和实际搜索不匹配。

【需求描述】

对于简拼搜索（组合搜索不支持简拼），当返回多个不同的纠错词时，只取热度最大的纠错词对应的结果，如果该纠错词对应多个类型，则各个类型都返回，按热度降序。如：MBY还有一个纠错结果是“毛不易”影视名，则搜索“MBY”时，纠错接口返回“歌手毛不易”和“影视名毛不易”两个，歌曲“马步谣”、专辑“梦，不远”则过滤掉。

**6.3优化组合搜索存在单字时的纠错效果**

【需求描述】

当前：很多单字不入纠错词典，导致单字组合搜索时经常会出现错误纠错结果。要求单字也正常入纠错词典，若组合纠错结果与其他类型纠错词相同，则优先取其他类型纠错词。如搜索“爱陈慧琳”，满足组合有关联“歌手陈慧琳 歌曲爱”，若同时有个歌曲就是“爱陈慧琳”，则优先取“歌曲 爱陈慧琳”的结果。

badcase:“爱 陈慧琳”、“飞 柯有伦”等

**7.数据处理**

**7.1 MV运营播放量同步**

【需求描述】

原有：因数据源不同，导致同步过来的的播放量数据和音乐app上的播放量数据对不上；

改为：直接取音乐运营播放量，使得双方保持一致。

**7.2 24bit至臻音质处理**

【需求描述】

对于满足条件的歌曲增加24bit的标识，条件如下：全曲下有如下格式码**011005、011006、011007。**

**7.3 action新增状态处理**

【需求描述】

曲库接收中音数字专辑接口、分发门户数字专辑接口，action字段增加 5. 隐藏 6. 恢复。对于隐藏的数字专辑，搜索结果中需要过滤。

**7.4 产品判断改用status**

【需求描述】

原有：根据中音分发的产品操作类型action字段来判断产品状态；

改为：曲库各产品接口增加产品状态字段status，搜索此字段来识别产品状态。

# V2.11.0

**1.音乐app支持语义搜索 – 280 遗留**

【需求描述】

当搜索词包含汉字，且搜索词长度大于等于4时，可以判断是普通搜索词还是语义搜索词，可以从语义搜索词中提取关键词以及关键词类别。

1. 搜索词包括单个关键词

* 单个关键词：歌曲名、歌手名、词曲作者、专辑名、影视/综艺名、标签名~~、歌词~~；

1. 组合关键词

* 组合关键词1：歌手名+歌曲名、歌手名+专辑名、歌手名+综艺名、歌手名+标签名、标签名+标签名；
* 组合关键词2：单个关键词+排序、组合关键词+排序。

1. 支持语义同义词

支持在运营管理平台配置语义同义词，搜索接口支持去运营配置，如：

* 给标签“钢琴”配置同义词“钢琴曲”，可通过“钢琴曲”搜索出带有“钢琴”标签的歌曲；
* 给“舒缓”配置“驰放|抒情流行”的同义词，可搜索出带有“驰放”或者“抒情流行”标签的歌曲。

\*注：

最佳展示位、无维度、歌曲、专辑、视频、~~歌词~~均需要支持语义搜索。

**【举例】**

1. 单个关键词：

* 输入“我想听花香”，输出关键词：花香，关键词类别：歌曲；
* 输入“推荐歌手李白的歌曲”，输出关键词：李白，关键词类别：歌手；
* 输入“我想听钢琴曲”，输出关键词：钢琴，关键词类别：标签。

1. 组合关键词1：

* 输入“我想听周杰伦的告白气球”，输出关键词：周杰伦，类别：歌手；关键词：告白气球，类别：歌曲；
* 输入“播放陈奕迅的粤语歌”，输出关键词：陈奕迅，类别：歌手；关键词：粤语，类别：标签。

1. 组合关键词2：

* 输入“周杰伦的最新单曲”，输出关键词：周杰伦，类别：歌手；排序规则为按时间倒排。

输入“播放陈奕迅的最热粤语歌曲”，输出关键词：陈奕迅，类别：歌手；关键词：粤语，类别：标签；排序规则为按热度倒排。

**2.票务独立搜索**

【需求描述】

咪咕购票平台页面，增加独立的票务搜索。增加独立票务搜索接口，支持搜索出票务数据，搜索无结果需要支持返回无结果推荐。运营管理平台需支持票务置顶配置、查看（类型、来源、演出时间、上下架状态筛选，支持票务名称、艺人名称、城市、地点查询，支持双击id查看票务详情）、上下架、无结果配置。



**【参与搜索字段】**

场次名称、演出城市、演出地点、演出分类、艺人名称；

**~~【参与筛选的字段】~~**

~~演出城市、票务类型（选座、不选座）、演出分类；~~

**【排序规则】**

1. 支持人工置顶。当搜索词与运营配置搜索词精确匹配时，优先返回人工配置的内容；
2. 精确匹配 > 前中缀匹配 > 模糊匹配；
3. 演唱会>音乐会>话剧舞台剧>亲子；
4. 前两者相同情况下按艺人热度排序；相同艺人且同匹配度情况下，按演出时间排序，优先展示最近场次；

**【其他规则】**

1、搜索平台需返回票务场次id和来源（00：自有；01：永乐）；

2、已下架项目不支持搜索；

3、场次名称、演出城市地点、艺人名称不重复加分。

**3.支持客户端扩展活动/专区扩展跳转方式**

【需求背景】

原有：仅支持跳转h5页面，且页面统一显示分类名称为“专题”。



需增加：

* 支持路由跳转（歌单分类、榜单）
* 活动/专区支持关联内容id（点击跳转内容列表页面）

【需求描述】

1. 对于运营管理平台

* 支持配置专题活动所属分类。支持在专题、活动、榜单、专区 4个分类中进行筛选，默认为专题；
* 支持增加活动/专区的跳转方式。包括：1）支持运营pageid，前端通过pageid跳转音乐app的原生界面（单值）。2）支持配置关联的id（支持多值），用户在前端点击后，根据id跳转列表页面。（当前仅需要支持横屏视频彩铃、竖屏视频彩铃、歌单即可）。

1. 对于搜索平台，支持返回运营管理平台的配置信息，以便于前端进行展示。
2. 对于门户。需要返回前端展示的各种信息。对于id跳转列表页，需和前端讨论是否要由门户返回封面、名称等详情。

**4.支持返回搜索倾向**

【需求描述】

1、对运营管理平台。支持配置某个搜索词的搜索倾向（歌曲、专辑、视频、歌单、歌词），在运营管理平台配置某个词的搜索倾向，搜索平台根据运营配置的倾向性搜索对应的类型，搜索倾向默认是歌曲。

**5.支持标签自动展示**

【需求描述】

当满足以下任何一种情况，自动展示标签内容：

* 当普通搜索无结果，标签有结果时，自动将标签内容填充到搜索结果列表中（当前已实现）；
* 当命中同名歌曲时（包括歌手名/专辑名匹配），判断该标签的标签强制展示是否选中，若选中，自动展示标签列表，否则展示歌曲结果，待用户点击时查看时才展示标签列表内容。

1. 对于运营管理平台。在内容管理——》标签歌单菜单，增加标签强制展示的配置项，默认选中，支持取消选中，支持批量操作选中状态。



**6.歌曲列表首位支持展示所属榜单/分类**

【需求描述】

1. 爬虫。自动爬取咪咕音乐上的榜单，将榜单中歌曲、歌曲对应的排名以及榜单对应的h5/pageid取出。若一首歌曲存在于多个榜单，则对比歌曲在各个榜单中的排序，优先取排名较高的那个。
2. 搜索接口。将对应的排行榜名称+排名拼接成 所属歌曲的榜单描述，并返回榜单对应的 url/pageid，用于前端在点击该描述后可以跳转榜单详情页。

注：

根据咪咕不同榜单的更新机制不一样，需要根据榜单的更新时间定时刷新。

**7.数字专辑排序和展示规则调整**

### **7.1 数字专辑排序规则调整**

【需求描述】

原有：在最佳展示位、专辑tab下、歌手下专辑排序，数字专辑均享受最高的优先级；

改为：仅在销售期的数字专辑享有该优先级。

### **7.2 销售标识判断调整**

【需求描述】

原有：仅对于首发期的数字专辑展示“数字专辑销售中”的标识；

改为：只要专辑未超过销售截止日期，均可展示上述标识；

**8.歌曲标题简化（标题+专辑名+描述去重）**

【需求描述】

若满足以下两种条件中的任何一个，在搜索结果显示歌曲列表时，过滤歌曲名称尾部的括号和括号内容:

* 歌曲名称尾部括号中的文字与包含于专辑名或者歌曲描述中；
* 若歌曲名称尾部括号内的《》中的内容包含于专辑名或者歌曲描述中。

若满足以下条件，则在搜索结果的歌曲列表中不显示歌曲描述：

* 描述中《》中的内容包含于专辑名中。

badcase：

如下，根据上述规则过滤之后，两处红框的内容均不应该显示。多见于综艺和影视歌曲。



**9.视频彩铃搜索优化**

### **9.1彩铃搜索规则优化**

【需求描述】

当前：在综合搜索入口，当用户搜索词中包含“视频彩铃”、“彩铃”等配置内容时，最佳展示位支持将这些字样过滤后匹配彩铃和对应的歌手名。但是，当用户点击查看更多进入彩铃搜索入口时，由于不支持，会导致无对应结果。

改为：在彩铃搜索入口搜索时，增加支持如下场景，以和综合入口的最佳展示位保持统一：

* 彩铃名称+彩铃字样：“说爱你视频彩铃”；
* 歌手名称+彩铃字样：“周杰伦的视频彩铃”；
* 歌手名称+彩铃名称组合：“周杰伦的视频彩铃说好不哭”；
* 彩铃字样+彩铃名称：“视频彩铃小精灵TFBOYS”。

同时，原有的模糊匹配规则保留。

Badcase：

### **9.2放开type限制**

【需求描述】

原有：彩铃搜索入口，仅支持type = 0的视频彩铃类型；

改为：彩铃搜索入口，仅屏蔽 type = 4 的视频彩铃类型（UGC）。且匹配度相同时增加如下优先级：

1、包月库+文件格式9:16；2、包月库+文件格式4:3；3、基础库+文件格式9:16；4、基础库+文件格式4:3，相同类型下按热度排。

## 10.增加特殊词库配置，特殊词库内容不分词

【需求描述】

运营管理平台增加特殊词库配置，当搜索词与配置中的某一项精确匹配时，该词不进行分词匹配。

对于运营管理平台：1、支持新增；2、支持模糊匹配已配置的特殊词；3、支持修改/删除已配置的特殊词。

badcase： 搜索“抖音”，排名靠前的歌单里有“电子电音抖腿大法”这类分词匹配的。

## 11.联想提示单字母的缓存机制

【需求描述】

将26个字母的联想提示结果做缓存，缓存结果定时刷新，并支持运营可配置。

## 12.搜索置顶/标签过滤无版权内容

【需求描述】

搜索接口返回置顶和标签列表内容时，自动过滤无版权内容。

## 13.纠错优化

【需求描述】

1. 单一纠错和组合无关联纠错优先级调整。

原有：纠错时，组合无关联默认优先于多字/错字；

改为：只有当组合无关联为歌手+歌手、歌手+歌曲、歌手+专辑时，优先级才高于多字/错字。

badcase： 小娃娃脸，纠错为“小娃娃”和“脸”，未纠错为“娃娃脸”。

2、取消纠错的正则表达式，改为由语义去实现（需要罗列下哪些可以由语义支持

1. 纠错词典中区分别名和原名

## 14.搜索排序优化

### **综合搜索入口**

【需求描述】

1、666分组增加过滤条件。对于非热门的666分组（普通歌曲>普通歌手>普通专辑），必须保证组内的歌曲热度hot > 0或者hotlevel >0，否则，该歌曲的分组优先级取消。

badcase: Justin Bieber。

2、最大概率匹配。对于通过歌手/专辑/影视名匹配的结果，若同时匹配了歌曲名称，则该歌曲排在对应匹配结果最前面。

badcase: 家III，同名歌曲没有排在前面。

3、模糊搜索的热门内容支持优先展示

当前：精确匹配组优于前缀匹配组和分词模糊搜索匹配组，导致部分热门内容排序靠后；

改为：增加两个最佳匹配组，当精确匹配组最高热度为0且hotlevel<1时，从满足如下条件的结果里取两个排在精确匹配组前面：

1. 通过前缀或分词模糊（要求匹配度大于等于75%）匹配；
2. 歌曲hotlevel > 0或者hot > 50万；
3. 有版权。

badcase： 同一个世界。黄绮珊和萧敬腾的“当我们在同一个世界”应该排在首位。

### **演唱会搜索入口 -- 优先级高**

【需求描述】

演唱会支持中缀匹配，前中缀优先级相同。Badcase：十一，搜不到第十一届咪咕汇的相关演唱会资源。

## 15.运营管理平台

### **1.1新埋点统计数据展示**

【需求描述】

1. 搜索量、转化率统计。
2. 搜索量&转化率 菜单，数据来源改为music\_new目录下的文件：



1. 统计每天不同来源的搜索量占比和转化率，数据来源为music\_new目录下的文件：



1. 搜索点击详情事件展示，数据来源为music\_new目录下的文件：



支持根据列1精确/模糊搜索、支持根据列 2、3进行筛选，支持根据时间段进行筛选。



1. 支持展示特殊展示位的点击量，数据来源为music\_new目录下的



6.支持根据列0中的id精确匹配，支持根据列1中的名称模糊匹配；支持根据列、2、3进行筛选，支持根据时间段筛选（一个月）。



### **1.2视频彩铃查看增加筛选条件**

【需求描述】

数据查看——>视频彩铃菜单。

1. 增加视频彩铃版权状态筛选；
2. 增加视频彩铃类型筛选。

### **1.3 标签歌单支持模糊匹配**

【需求描述】

数据查看——>标签歌单。支持标签歌单模糊匹配，支持运营在模糊搜索和精确搜索间进行切换。

### **1.4 数据查看支持去符号匹配**

【需求描述】

当前搜索词中包含空格或者特殊符号，匹配不出来对应的内容，需支持去符号匹配。

1. 搜索词带符号，匹配内容不带符号，可检索出来；
2. 搜索词不带符号，匹配内容带符号，也可以检索出来。

### **1.5 最佳展示位干预支持根据干预类型进行筛选**

【需求描述】

推荐管理——>最佳展示位干预，增加一个干预类型（类别）的筛选，支持进行更新时间的筛选（精确到天）。

## 16.数据处理

### **1.1 票务数据同步**

【需求描述】

由搜索平台提供接口，音乐票务平台调用该接口自动同步数据，同步机制：

* 每天全量同步一份票务数据（包括自有和第三方）；
* 自有的增量数据会实时同步，第三方的增量数据无单独的实时同步，跟随每天的全量同步进行更新。

具体参数参见接口文档中的3.5.1.2。

### **1.2 抽屉中后同步的录音棚版在抽屉中的排序优化**

【需求背景】

对于某一热门歌曲，在当前无录音棚版时，运营可能会增加live版等的hotlevel值，但是后续若在增加了录音棚版时未及时运营，会导致抽屉排序时，live版的优先于录音棚版的展示。

【需求描述】

当新增录音棚版内容同步时，将live版的hotlevel可以自动更新给录音棚版。

### **1.3 专辑/MV热度更新优化**

当前：专辑/MV热度都是取自对应的歌曲，当歌曲热度调整后，对应的专辑和MV的热度并未同步更新；

改为：增加定时任务更新专辑/MV的热度。

### **1.4 版权判断规则变更**

当前：歌曲版权计算考虑全曲和彩铃。

改为：只考虑全曲的版权有效期，如果全曲无版权，则认为该歌曲无版权。

# V3.0.0

1. **中缀匹配实现方案调整**

【需求描述】

当前：中缀匹配时通过将搜索词转拼音匹配的，存在误搜漏搜；

改为：通过单字分词匹配。

【涉及内容】

1. 最佳展示位热门歌手、专辑 。逻辑与之前一致，对于热度大于20万或者hotlevel>0的支持前中缀，其他时候还是精确匹配。
2. 最佳展示位的期刊tab。中缀匹配优先于分词匹配。
3. 歌曲tab。中缀匹配优先于分词匹配。
4. MV tab。中缀匹配优先于分词匹配。
5. 专辑tab。中缀匹配优先于分词匹配。

badcase：

1、歌曲搜索“环环”搜不到“世界与你环环相扣”;

2、最佳展示位搜索“swift”,最佳展示位未匹配到Taylor swift （测试时注意取消置顶配置） 。

1. **歌曲搜索取消热门分词结果优先**

【需求描述】

当前：前缀匹配或者分词匹配75%以上的热门歌曲会调整到热度为0的精确匹配结果前面；

改为：上述前缀匹配的逻辑保留；分词匹配的逻辑取消，改为中缀匹配；增加当长度大于2时，如果错一个字，也支持将该结果调整到精确匹配的热度为0 的结果之前。

PS：原有最多只调整2个热门的在热度为0的精确结果之前的逻辑依然保留。

Badcase：爱情的迫降

1. **歌曲搜索、标签搜索支持语义**

【需求描述】

当搜索词包含汉字，且搜索词长度大于等于5时，可以判断是普通搜索词还是语义搜索词，可以从语义搜索词中提取关键词以及关键词类别。

1. 搜索词包括单个关键词

* 单个关键词：歌曲名、歌手名、词曲作者、专辑名、影视/综艺名、标签名~~、歌词~~；

1. 组合关键词

* 组合关键词1：歌手名+歌曲名、歌手名+专辑名、歌手名+综艺名、歌手名+标签名、标签名+标签名；
* 组合关键词2：单个关键词+排序、组合关键词+排序。

1. 支持语义同义词

支持在运营管理平台配置语义同义词，搜索接口支持去运营配置，如：

* 给标签“钢琴”配置同义词“钢琴曲”，可通过“钢琴曲”搜索出带有“钢琴”标签的歌曲；
* 给“舒缓”配置“驰放|抒情流行”的同义词，可搜索出带有“驰放”或者“抒情流行”标签的歌曲。

**【举例】**

1. 单个关键词：

* 输入“我想听花香”，输出关键词：花香，关键词类别：歌曲；
* 输入“推荐歌手李白的歌曲”，输出关键词：李白，关键词类别：歌手；
* 输入“我想听钢琴曲”，输出关键词：钢琴，关键词类别：标签。

1. 组合关键词1：

* 输入“我想听周杰伦的告白气球”，输出关键词：周杰伦，类别：歌手；关键词：告白气球，类别：歌曲；
* 输入“播放陈奕迅的粤语歌”，输出关键词：陈奕迅，类别：歌手；关键词：粤语，类别：标签。

1. 组合关键词2：

* 输入“周杰伦的最新单曲”，输出关键词：周杰伦，类别：歌手；排序规则为按时间倒排。
* 输入“播放陈奕迅的最热粤语歌曲”，输出关键词：陈奕迅，类别：歌手；关键词：粤语，类别：标签；排序规则为按热度倒排。

1. **高亮词优化**

【需求描述】

针对搜索接口和联想接口的返回的高亮词进行优化。

原有：原词及纠错结果作为高亮词返回。

改为：在原有的基础上，对于中文搜索词：对比搜索词与搜索结果，按字符比较相同部分也作为高亮词返回。如结果”世界美好与你环环相扣“， 搜索词”世界环环美好“， 会返回原有高亮词外，还会返回“世界美好”、“环环”。

其他case：10送红军。

1. **纠错功能**

**5.1英文纠错优化（待定，需要高岩峰再确认下）**

【需求描述】

英文纠错，支持所有单词都有错误的的纠错。每个单词错误的字母数，不多于四分之一，其中不足一个字母的，支持错一个字母（即3\*1/4 = 0.75，支持错一个， 5\*1/4 = 1.25，还是只能错一个）。不支持一个字母单词的纠错。

**5.2 纠错优化**

【需求背景】

解决小哑巴误纠为小鸭子的问题。

badcase：搜小哑巴，未能检索出专辑：哑巴，纠错返回结果为：小鸭子

**5.3性能优化**

1. **门户**

**6.1 歌曲名简化的优化**

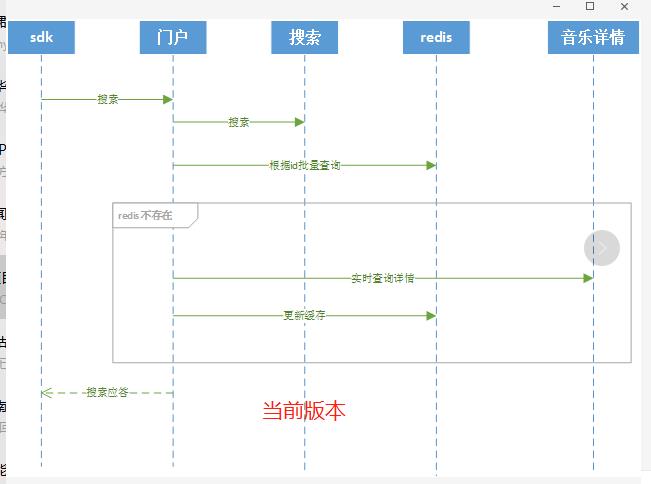
【问题处理】

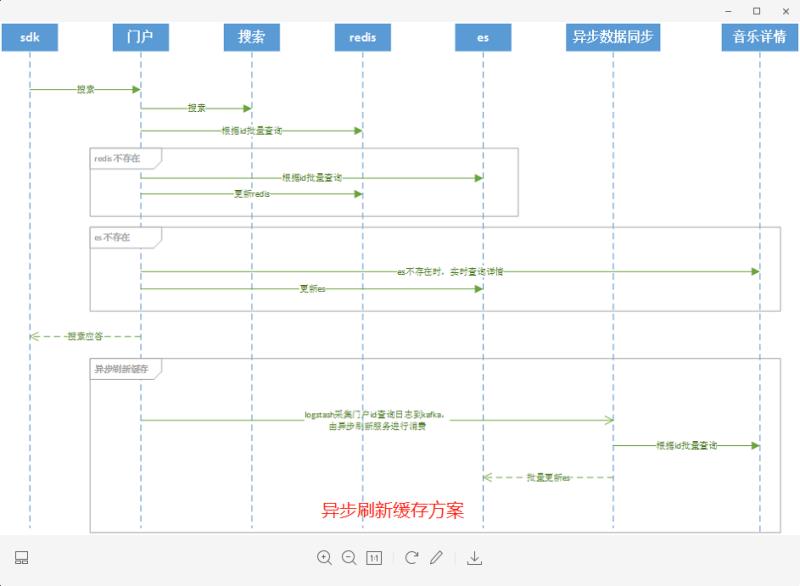
搜索英雄本色，歌曲名称尾部括号内的《》中的内容包含于专辑名或者歌曲描述中，未过滤歌曲名称尾部的括号和括号内容。

**6.2 数据缓存改造**

【需求描述】

对于依赖音乐侧的数据缓存，异步查询保存到门户侧。





上图中 es改为mongo

建议对门户接口性能进行压测。

# V3.1.0

## 1.支持前端的若干改造 -- 详见客户端v3.0.0prd

【需求描述】

1、最佳展示位支持同区内容同时展示；

2、支持返回搜索点击比例；

3、新歌预约。

## 基于用户行为的自动排序

【需求描述】

1. 数据来源：根据搜索sdk上传的埋点数据，由amber统计用户搜索后点击了哪些内容，文件参考需求1中的点击率；
2. 排序规则：

先只在分组（666分组）内进行排序调优，各分组顺序不变。组内排序改为：优先排isnew=1的，然后hotlevel=6的，再按照C倒序排列，最后按照hot倒叙排。

A：前日的搜索点击量；

B：昨日的搜索点击量；

点击意愿C=（A+B）\*（B/A）

【运营管理平台】

支持点击意愿数据的修改。

badcase：搜索罗大佑，当前第一位的“请珍重”，无多少点击量，光阴的故事点击量基本都在第一位。

## 3.联想提示优化

1、联想词来源

1. ~~单字母的自动更新~~

~~将26个字母的联想提示结果做缓存，缓存结果定时刷新，并支持运营可配置。建议直接复用联想置顶的redis，可以减少额外的开发；~~

1. 联想词库
2. top10000热搜词选择

* 改用新文件；
* 之前只要求热搜词必须匹配歌手名/歌曲名/专辑名/才能入联想库，现在增加匹配影视名称、支持匹配歌手歌曲专辑的匹配别名译名。
* 原有要求ophot设置为6，改为100。

1. 运营热词导入

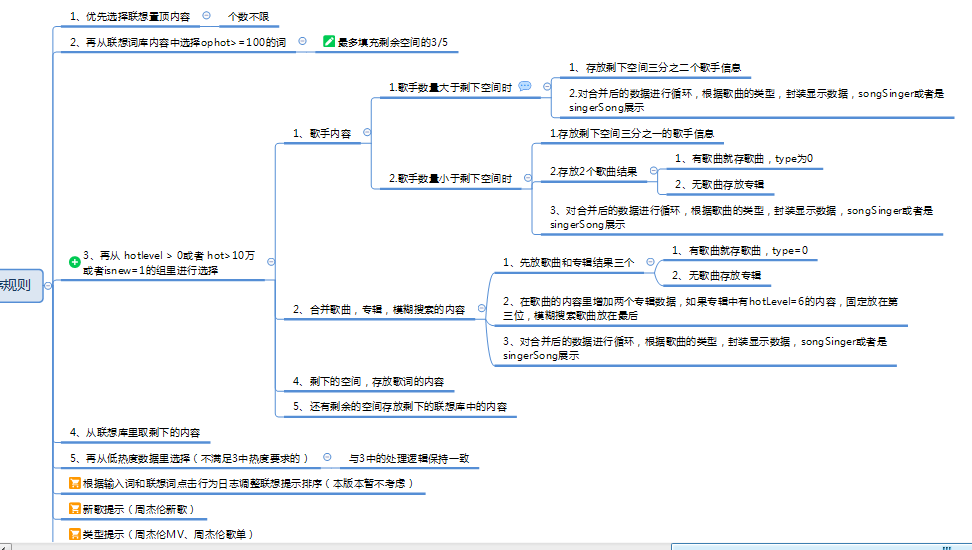
每日8点、12点、18点定时拉取音乐的热词（只取前6个），更新到联想词库，hot设置为10万，hotlevel设置为100。

1. 其他的参考下图中的要求即可（绿色logo标识的表示新增或修改的）



1. 联想词排序

（黄色标识的是后面版本要弄的，这个版本先不用考虑）



## 歌手下歌曲排序打乱

【需求描述】

歌手下歌曲还是按照热度排序，但是增加一条附加条件：属于同一抽屉的内容不能连续展示在同一页（总数少于1页的情况除外），抽屉内其他版本随机插入到后续其他页码中。



## 专辑isnew标识参与排序

【需求描述】

1. 最佳展示位。优先级：在售数字专辑>歌手>普通专辑和非在售数字专辑>….。若同时匹配多个在售数字专辑，优先返回带isnew标识的；同样的，若果同时匹配多个其他专辑，优先返回带isnew标识的。
2. 专辑tab。同上。
3. SAM专辑。不用改。
4. 联想提示。

## 歌单搜索优化

【需求描述】

当歌单内有版权歌曲数少于总数的1/3时，在专辑（无维度和SAM）和歌单排序时做降权处理：

有版权歌曲数超过1/3的用户歌单优先于少于1/3的小编歌单。



## ~~智能屏蔽~~

~~【需求背景】~~

~~当前音乐搜索的屏蔽只有一种，即当搜索词与屏蔽词完全相同时，联想提示和搜索均不返回结果。但是在很多场景下这样一刀切会导致屏蔽很多优质内容。~~

~~【需求描述】~~

~~将屏蔽词进行分类， 根据不同类型做处理：~~

~~A类（屏蔽完整搜索词）： 即当前的处理方式。用户无法通过屏蔽词精确搜索/联想到任何内容，包括歌曲、专辑、视频、歌单等。但本身内容不做屏蔽处理。~~

~~B类（屏蔽所有符合屏蔽词名称的歌曲/视频）：即：凡是歌手/歌曲/专辑/视频(包括视频彩铃)/歌单/歌词 = 屏蔽词的内容资源，均需要进行搜索屏蔽~~

~~【运营管理平台】~~

~~屏蔽词配置的时候支持设置类型，默认为A类。~~

## 无结果推荐问题处理

【需求背景】

若某一次请求超时导致搜索无结果，会返回无结果推荐内容。由于无结果推荐的翻页查询和普通搜索的翻页查询相同，若在无结果推荐内容点击进行翻页的时候又无超时了，又会展示正常的搜索结果，而不是无结果推荐的翻页查询结果。

【需求描述】

超时导致无结果时给前端单独的错误码用于标识该场景，前端后续提示用户重试。不直接返回无结果推荐内容。

若后台无法区分是超时还是其他原因导致的无结果，就一次性返回所有无结果推荐内容（限制最多返回30个）。

## 支持手动清除缓存

【需求描述】

1. 视频彩铃的缓存时间改为2分钟，当前为5分钟；
2. 支持视频彩铃、无维度、歌曲、专辑、视频、联想提示接口支持手动清除某个/些搜索词的缓存，使得可以搜索到最新的结果，提高重点保障场景下的搜索实时性。

【运营管理平台】

看是在运营平台上增加配置还是直接跑脚本处理。

## 抽屉排序优化

【需求描述】

1、抽屉过滤

当前：歌曲搜索结果中，如果同名歌手的同名歌曲会被放在一个抽屉里；

改为：歌手同名歌曲同名时增加如下判断

1）不同语言的不放在同一个抽屉里

case：谭咏麟的水中花存在国语和粤语两个版本，需要将两种语言的分别放在不同的抽屉中展示；

2）有新歌标识的不放在抽屉里

case：taylor swift 的 the man 新出了一个现场版，在新歌标识生效期间，需要将该版本在抽屉外单独露出，可方便排序控制。

2、运营热度自动同步。当新增录音棚版内容同步时，将live版/其他录音棚版的hotlevel可以自动更新给新同步的录音棚版。

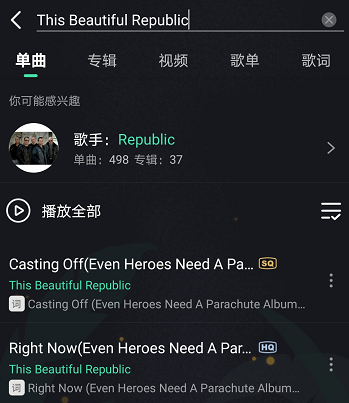
## 歌曲搜索歌词匹配优化

【需求描述】

当前：歌曲tab下，当搜索结果是通过歌词匹配出来时会返回三行歌词；

改为：

增加判断：若搜索词包含与歌手/专辑名中（去符号匹配），不返回歌词。case如下：

## 音箱搜索优化

【需求描述】

1. 支持分页。当前是将所有结果一次性返回，容易导致前端内存溢出，现改为根据前端传参控制每次返回的数量。
2. 组合搜索支持标签同义词。当前仅支持单一标签搜索时匹配标签别名，现要求组合搜索时也支持标签别名，如小星星的标签为“儿童歌曲”，若“儿童歌曲”标签配置了别名“儿歌”，搜索“小星星 儿歌”，也可以召回正确的结果。
3. 组合搜索精确匹配时支持返回多值。针对歌曲名+标签组合搜索，当精确匹配多个结果时，支持将所有结果都返回。（当前只返回第一个）

badcase：小星星，儿歌

## 爬虫

### **歌手热度计算元数据爬取支持匹配别名**

【需求背景】

如celine dion在QQ音乐的歌手库原名是带重音符号的，当前仅用原名与QQ音乐的歌手原名比较，无法匹配上，元数据爬取不了就计算不了热度。

【需求描述】

1、爬取元数据时，支持将咪咕音乐的歌手原名和别名与QQ榜单上的原名/别名匹配；

2、从QQ歌手上述热门歌手需要 *每日* 更新热度。

### **热门歌手的歌曲热度支持定时自动更新**

【需求描述】

对于hot大于5万或者hotlevel >1的歌手，先全量更新一下这些歌手的歌曲热度，对于本次热度大于10万的歌曲，后续 *每周* 都要更新下歌曲热度，如果有降低到10万以下的歌曲，后续可不再更新，自行衰减热度即可。

注：歌手和歌曲热度计算时需要保存计算热度的元数据（播放量、评论数之类的），支持在页面上查看这些元数据。

## 纠错优化

【需求描述】

1. 纠错词典中区分别名和原名

具体要解决啥问题给忘记了，纠错这边先给区分开，等后面看下搜索接口咋根据不同类型来处理。

1. “仔很忙 周杰伦”这类纠错问题处理

待定。

## 数据处理- MV关联歌曲时增加歌曲状态判断

【需求描述】

分发MV内容时，如果有关联歌曲id，需要对歌曲id的状态做判断，如果该歌曲素材没有分发给搜索，那么清除MV关联的歌曲id。

如果后续歌曲分发过来了，能不能再跟MV关联上？

# V3.1.1

## 问题需求补丁描述：

### **1. 歌曲与专辑同名的优先顺序与用户意愿排序冲突，导致同名的提到最前面了 -- 00750091655**

修复：只在相对应的分组内优先

示例：搜索“再回首”， 专辑名与歌曲名相同的提到了最前面，应该只提到专辑匹配的组的最前面。

### **2. 修改静态检查错误，导致分组内hotLevel比较失效 -- ? 00750091657**

查其他问题时，发现修改静态检查把分组内排序逻辑改错了。导致分组内排序没有完全按照 isNew > hotLevel=6 > 点击意愿 > hotLevel > hot的规则。验证时可按此规则验证。

### **3. 歌词搜索匹配的，如果专辑名或歌手名包含搜索词则去掉歌词显示，但对于英文包含关系只做了严格比较 -- 00750091658**

修复：优化为不区分大小写

示例：搜索 "this beautiful republic"，首字母小写的时候仍然把歌词显示出来了，不区分大小写更合适。

# V3.1.2

## 支持5G专区的搜索

【需求背景】

咪咕音乐上有个5G专区，里面分为若干子专区，其中全面屏和24bit专区需要提供搜索功能。前后端由我们负责，音乐专区h5页面增加搜索入口，点击跳转我们的原生搜索页面。

### 支持单曲24bit专区搜索

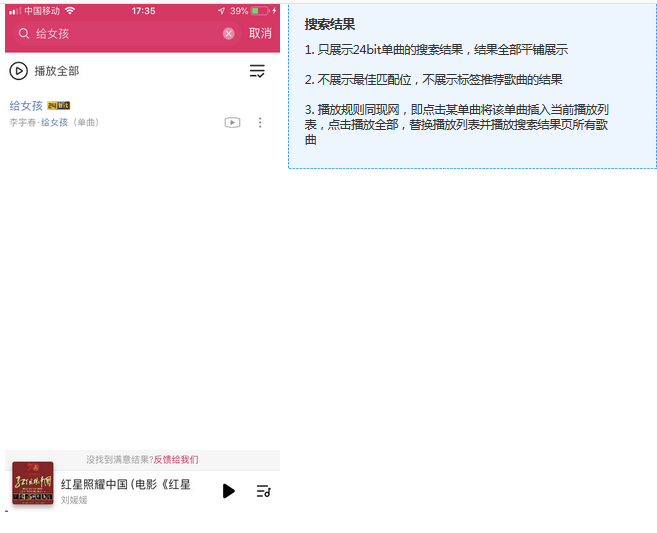
【需求描述】

1. 对于平台：支持联想提示和歌曲搜索。要求：

* 联想和搜索仅支持24bit音质的歌曲（如果24bit音质对应的产品无版权则不用召回）；
* 采用打乱排序（结合前端传参）。

1. 对于前端：

* 搜索启动页面：仅包含搜索历史记录（与综合搜索分开），不需要调用接口获取默认搜索词；
* 输入过程中：支持联想提示；
* 搜索结果页：仅支持展示歌曲，不支持展示其他tab内容和最佳匹配位



### 支持全面屏视频彩铃搜索

【需求描述】

1. 对于平台。仅支持全面屏视频彩铃（9:16）的搜索和提示；
2. 对于前端：

* 搜索启动页面：仅包含搜索历史记录（与综合搜索分开），不需要调用接口获取默认搜索词；
* 输入过程中：支持联想提示；
* 搜索结果页：仅支持展示视频彩铃，不支持展示音频彩铃tab，点击封面跳转视频彩铃播放页。

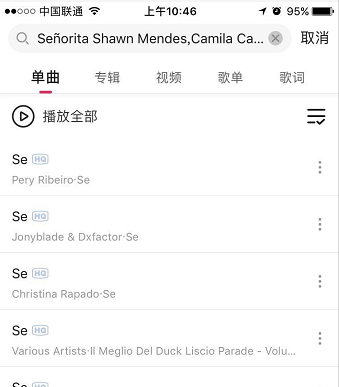


## 语义搜索减少无匹配场景优化

【需求描述】

1、英文输入比较长时误匹配场景优化，先强制限制搜索词超过20个词不支持语义搜索，后续讨论优化方案。

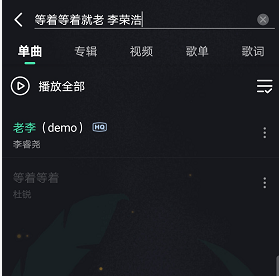
badcase:"Señorita Shawn Mendes,Camila Cabello"



2、取消歌曲+歌曲，歌曲+专辑组合场景

    由于用户实际使用过程中极少有歌曲+歌曲组合，歌曲+专辑组合场景，所以语义取消支持此两种场景，减少误匹配概率。

badcase：“等着等着就老 李荣浩“，误识别为“等着等着”和“老李”了



## 用户热搜top10000入联想词库支持匹配歌手+歌曲组合

【需求背景】

          目前会根据用户的搜索行为记录优化联想提示。要求在用户热搜top10000的词与搜索库中匹配且转化率满足要求的可以入联想提示库，优先展示。目前只匹配了歌曲、歌手、专辑库，没有支持匹配歌手+歌曲，会导致一部分热搜的组合词无法优先提示。

## 音乐可视化上报规则变更

【需求描述】

            除了跟歌曲、歌手、专辑库里内容匹配，也需要和歌手+歌曲/歌曲+歌手匹配，保证热搜的组合词可以优先提示。

【需求描述】

          原有：上报所有产品同步日志；

          改为：同个产品，只上报最新的一条。

# V3.1.3

## 1.切换资源池 redis使用命令改造

【需求描述】

资源池redis 禁用keys命令，现在代码码中使用了此命令，切资源池时代码需要配合整改。

【修改】

改为：scan方式遍历

## 2.缺陷问题修复

1. 【咪咕音乐】【v3.1.2】歌曲已删除，app上专辑tab内还是提示包含该删除歌曲--[00750093181](https://spms.migu.cn/spms/static/defect/index.html" \l "/defect/detail?defectId=93181&type=3)
2. 【咪咕音乐】【v3.1.2】联想提示出现联想词重复的问题-- 00750093323



1. 【咪咕音乐】【v3.1.2】联想接口，提示词热度高的排在后面 -- 00750093322



# V3.2.0

## 1.配合前端改造

### 1.1支持返回原唱描述和运营描述

【需求描述】

1. 描述分为两种：原唱描述、其他运营描述。
2. 对于运营管理平台，在内容管理里支持对歌曲描述增删改（包括描述名称和类型）；
3. 对于搜索平台和门户，支持返回运营平台配置的描述，且有如下要求：

* 由于前端展示时，词曲作者和歌曲描述在同一个展示位，要求词曲作者和歌曲描述不能同时返回。优先级如下：运营描述 > 词曲作者 > 原唱描述。

注：若命中了词曲作者，不展示歌词。

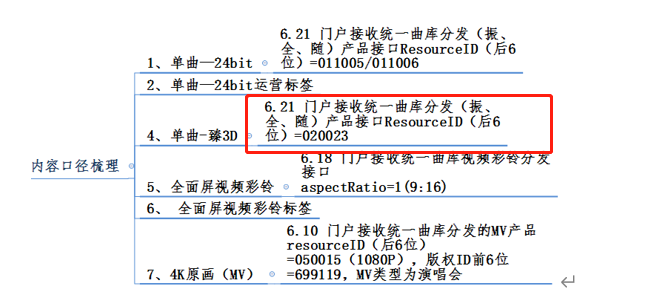
### 1.2支持返回3D标识（待确认）

【需求背景】

音乐会新增一部分3D歌曲。

【需求描述】

1. 对于数据处理。3D歌曲对应的振、全、随产品的Resourceid（后6位）都是020023。数据处理可根据该规则判断哪些是3D歌曲，给与3D标识。
2. 对于搜索平台和门户，支持返回该标识，前端展示在音质的同级位置上。



### 支持视频/歌单的筛选和排序控制

【需求描述】

1. 无维度搜索支持指定视频/歌单排序方式；

包括综合、最新、最热三种排序。

* 综合：默认选中状态，按照当前排序即可。
* 最新：对于视频，按照同步搜索时间倒排；对于歌单，按照强版权组》弱版权组排序，组内按照同步搜索的时间倒排。
* 最热：对于视频，可先按照评论数 》 播放量 》hotlevel\*10万+hot倒排（后续要考虑计算综合热度）。对于歌单，按照强版权组 》弱版权组排序，组内按照评论数 》 播放量 》hotlevel\*10万+hot倒排。

注：歌单强版权组指的是有版权歌曲数超过歌曲总数1/3的歌单，否则为弱版权组。

1. 无维度搜索支持按标签筛选
2. 支持返回当前歌单tab对应的标签列表。标签列表选择和排序的规则为：

* 标签选择：只将播放量大于0的，且有版权歌曲数占比超过1/3的歌单的标签提取出来；
* 标签排序：将上述歌单按照播放量倒排后逐个提取歌单所属标签。若标签多于10个，则只取前top10。

1. 用户点击标签后支持按照标签对当前歌单列表进行筛选 （支持多个标签组合筛选）。

### 热门搜索改造

【需求描述】

只涉及门户和运营平台。

新增一个热门搜索模块，该模块内容由搜索根据用户实际搜索行为自动生成（同时支持运营修改），暂定20个，展示内容包括：热搜词、热度值、角标：

1. 热搜词生成规则（自动）：

每个小时统计一次该时段内搜索量top100的热搜词，假设当前时段的搜索量为Ht，前一时间段的为Ht-1，则进行如下计算：

H = Ht (1+Ht/ (0.1\*Ht-1))。

按照H倒排选择top50的热搜词。然后增加如下判断：

* 过滤掉黑名单中的热搜词；
* 只保留热度最高的3个歌手类型热词，剩余的歌手类型过滤；（需要检查下现网统计的热词类型是否有问题，有的话还需要处理下）；
* 使用热搜词调用搜索接口，判断搜索结果第一位是否相同，如果相同，则只保留搜索量最高的；
* 查看是否有运营指定的热搜词及对应位置；
* 按照上述过滤后，取top 20的作为搜索热词填充的展示位中，多余的舍弃。

评审记录比如：运营指定位置135等，然后代码入的自动插进去 补全为20个；再比如说今天运营1、3 、5 明天修改为2、8、9等这个代码自动入夜不会去动这个。

1. 热度值计算规则：

* 展示热度（手动） = （前1热搜词热度+后1热搜词热度）/2 +rand (0~2000);用于在用户新增一个热搜词后，可自动更新该热搜词的热度值。
* 展示热度（自动） = 基础热度\*系数1 + 实际搜索数量\*系数2；

基础热度 = 暂定按照当前时段的搜索总量作为基础热度；

实际搜索量 = 当前时段的搜索量。

系数1和系数2的具体值待定。

1. 运营：
2. 支持管理热词、位置、角标；
3. 支持热搜词黑名单。在黑名单中的词不允许展示在热搜榜中。黑名单来源有两处：

* 运营手动添加。；
* 代码自动跑入（取热搜词top100中转化率小于70%的。每日更新）；
* 支持指定给某些用户才返回新样式的标识和热搜内容；
* 支持在后台统一关停该热搜模块（新增的这个模块）。

### 超多评论标识

【需求描述】

在非分词搜索结果列表中，若匹配项超过1条，找出评论数最多的那个歌曲，判断：

1. 该歌曲有版权；
2. 该歌曲评论数大于999；
3. 该歌曲无点击率标识（超过XXX点击）；

若上述三点均满足，则在原来的点击率标识字段中返回“超多评论”的文案。

另：运营平台支持增删改该文案。

注：后续根据播放情况可能还会增加“试听飙升”的文案，本版本暂不考虑。

### 5G专区视频彩铃搜索和提示调整

【需求描述】

原有：只支持竖屏彩铃搜索；

改为：横屏竖屏都支持，仅屏蔽type=4（UGC）的彩铃。

说明：排序规则按照现网即可，基础优先级如下：

1、包月库+文件格式9:16；

2、包月库+文件格式4:3；

3、基础库+文件格式9:16；

4、基础库+文件格式4:3；

相同类型下按热度排。

### 5G专区1080P 原画搜索

【需求描述】

要求搜索结果中过滤不是1080P的原画内容。

原画判断规则：cpid：699119，码率：050015的MV演唱会数据。

## 2.最佳展示位搜索优化

### 组合搜索优化

【需求描述】

1. 搜索词为 歌手+专辑/MV关联组合时，最佳展示位A区展示专辑/MV，最佳展示位B区展示歌手。注：此时若A区有运营干预项，则不展示歌手。

badcase： “我们的歌 王力宏”，需要把视频展示在最佳匹配位上。

### 前中缀匹配增加个数限制

【需求描述】

1、 对于原名。前中缀匹配时，要求搜索词字数占待匹配项字数的1/3以上（包括1/3）；汉字按照字数，英文按照字母数。

2、对于别名译名。由于在最佳匹配为上不展示别名译名，所以要增加限制。要求搜索词长度大于等于2时才能前缀中缀匹配别名，且要求搜索词字数占待匹配项字数的1/3以上（包括1/3）。

badcase： 搜索“菜”，最佳匹配位出现蔡徐坤；“同”搜索出萧敬腾。

### 支持歌曲关联歌单

【需求描述】

对热搜top 10000中入联想词库的搜索词，每天离线更新该搜索词对应的热门歌单。歌单选择方式：可以参考歌单搜索接口，选择最热排序，取top1的歌单。匹配位A区或者B区有空位时，可将该歌单置于对应的展示位上。

优先级：人工运营最佳展示位 》 按照搜索规则自动匹配的 》热门歌单。

限制：

1. 要求歌单是按照歌曲精确匹配，或者当搜索词长度大于2时，支持前中缀匹配（前中缀匹配优先级相同）。
2. 歌单播放量在50万以上；
3. 歌单内有版权歌曲数大于10，且有版权歌曲数占比不低于歌曲总数的2/3。

## 3.用户评论数干预搜索排序（联想暂不考虑）

【需求描述】

1. 热歌标识计算。若歌曲总评论数大于500且一周内歌曲评论数增量大于50，则给定热歌标识。
2. 对于666分组内排序

原有：优先排isnew=1的，然后hotlevel=6的，再按照点击意愿倒序排列，最后按照hotlevel > hot倒叙排。

改为：优先排isnew=1的——> hotlevel=6的——>有热歌标识的——>再按照点击意愿倒序排列——> 最后按照hotlevel\*10万+评论数\*100+hot。

## 4.增加数据露出

### 抽屉排序调整—仅针对咪咕音乐app

【需求描述】

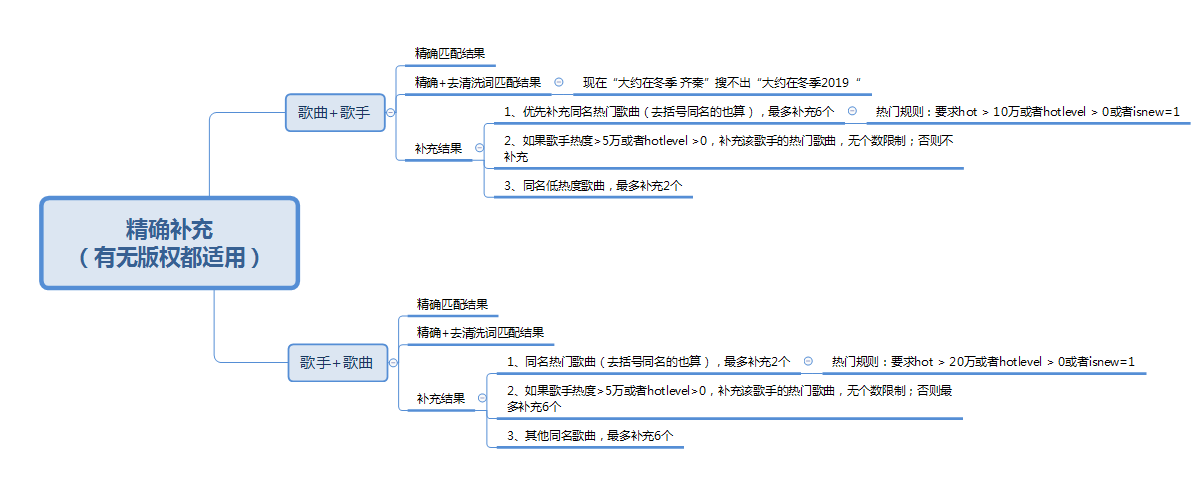
当搜索结果只有一个抽屉时（没有其他搜索结果）增加如下处理：将抽屉内有版权的内容全都露出，无版权的仍然折叠在抽屉中。

### 组合精确匹配时支持返回非组合结果—仅针对咪咕音乐app

【需求描述】

当前：“歌手+歌曲”精确匹配后，只返回精确匹配的歌曲结果；

改为：

**5. 联想提示1.2轮优化**

（第1轮是v3.1.0版本上线； 1.1 轮为top10000支持组合，v3.1.2版本上线）

【需求描述】

1. 问题处理。避免重复加分。Badcase：“最美”提示问题，“最美情侣”热度低于“最美的伤口”，但是提示在之前，是因为前缀和分词匹配重复加分导致的。



1. 关于过滤。此处指的是去除（）内容之后的长度。如果去除（）之后长度大于下述要求才过滤。

Badcase：“郭富城 浪漫樱花”提示不出来。可以多找一些综艺/晚会的歌曲测试。



1. 分组规则修改

之前歌手和歌曲统一限制的hot > 10万，现在改为歌手只要hot > 5万就可以放在热门组中参与匹配。



1. 模糊匹配场景

模糊匹配的不支持返回组合结果，只返回单一匹配结果。

badcase：“惊雷抖音版”。

## 歌曲名简化优化处理

【需求描述】

1. 支持去符号单字分词全匹配，

Badcase：亲爱的热爱的。



1. 过滤前做下待过滤部分的长度判断：

纯中文：3个字以上才过滤；

带字母：8个字以上才过滤。

## 高亮问题处理

【需求背景】

如下图所示，当前的高亮方案会导致搜索结果中高亮显示的非常乱。

【需求描述】

高亮词包括如下内容：

1. 搜索原词；
2. 纠错返回的实体词；
3. 人工纠错结果；
4. 语义识别实体词；
5. 分词结果：按字符比较相同部分作为高亮词返回，过滤单字的。

## 标签自动展示优化

【需求描述】

原有：当标签与歌曲/歌手/专辑同名时需要标签有强制展示标识才优先展示标签下内容，否则优先展示普通搜索内容。

改为：当普通搜索结果列表第一位无版权时，不管标签是否强制展示，都优先展示标签下歌曲。

## 语义搜索支持一词多义

【需求背景】

当前语义搜索单个实体仅支持单个属性，但是在实际数据中，一个词可能既是歌曲名，又是歌手和专辑名。

【需求描述】

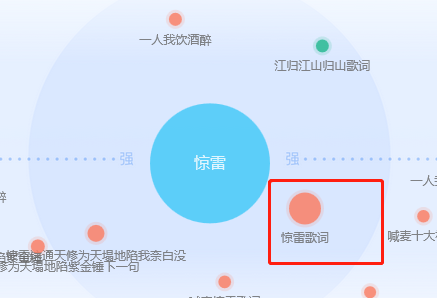
1. 单实体时，如果有多个属性，则在搜索时按照各个属性的热度进行加权。
2. 多实体时，如果有多个属性，则按照多个实体间的关联关系择优选择，如：“Dean的专辑beyond”，虽然beyond有歌手、歌曲、专辑多个属性，且歌手属性热度较高，但在刚刚这个场景，要优先选择专辑属性。

## 纠错需求

【需求描述】

1. 置顶/最佳展示位干预搜索词，只支持rel=1(组合关联纠错)的纠错；--
2. 标签名（标签歌单名称、精品标签名称）入纠错白名单；
3. 歌词纠错
4. 歌词入库条件：

* 选择hot>20万或者hotlevel > 0的歌曲；
* 用歌曲名称去百度指数上爬取关联词，如果关联词中有“歌曲名+歌词”，且关联度满足阈值（待根据爬取的数据确定），则将歌曲对应的歌词加入纠错词典；
* 需要过滤字数小于4的歌词。
* 入库时热度按照歌曲热度的1/10取值；



1. 之前遗留的两处需求

* 支持热门MV入库。热门标准：hot>10万或者hotlevel>0；
* 支持歌手/歌手别名+MV组合搜索。

## 数据处理

【需求描述】

1、过滤所有产品为空的歌曲（搜索不露出）；

2、过滤所有版权歌曲数为0的数字专辑（不包括最开始未分发歌曲的），主要有以下两种场景：

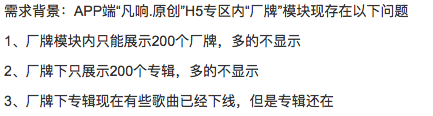
* 歌曲均被隐藏；
* 歌曲全为无版权的。

3、歌曲全为无版权的数字专辑不展示在售标识。

# V3.2.1

## 1.支撑音乐厂牌需求

【需求背景】



【需求描述】

1. 数据处理

需要同步厂牌的数据。所有的厂牌有一个共同的父栏目，每个厂牌对应一个子栏目，每个栏目信息里有这个厂牌关联的专辑id。

1. 搜索平台/门户

支持厂牌和厂牌下专辑的分页查询。无搜索词。排序规则如下：

1）厂牌查询排序

* 优先取厂牌榜单中的厂牌（在搜索运营平台配置）；
* 然后按照包含专辑数倒排，即包含专辑数最多的厂牌展示在最前面；

2）厂牌下专辑查询排序

* 优先取厂牌下专辑榜单列表内容（搜索运营平台配置）；
* 然后按照专辑的发布时间倒排，即最新发布的专辑优先展示。

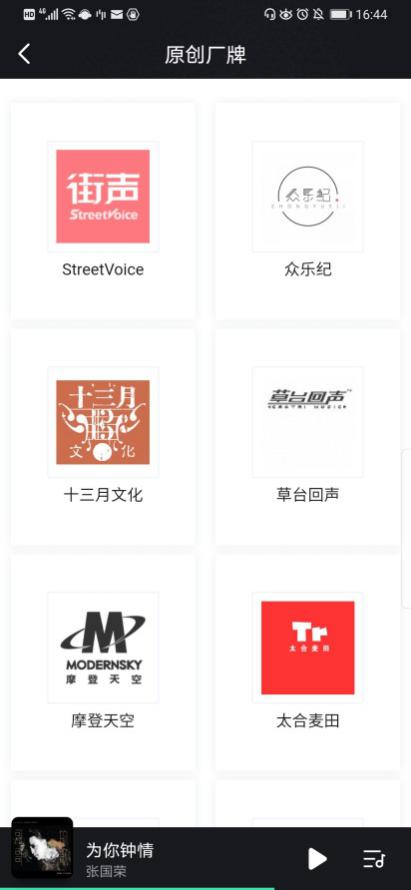
注：注意运营和自动查询的要去重。

1. 运营平台

* 支持厂牌榜单的运营。支持在文本框内输入一组厂牌id，排在前面的厂牌id在查询结果中优先展示。文本框内的内容支持清除。
* 支持厂牌内专辑榜单运营。类似于置顶功能，可以指定某个厂牌id下的专辑排序（普通置顶配置的是置顶搜索词）。

1. 查询入口说明

在【音乐】tab 拉到页面底端——>点击右侧箭头进入【凡响原创】专区——>拉到专区内页面底部。

## 2.彩铃专区优化

### 1.1联想提示优化

【需求描述】

1. 针对彩铃专区，支持视频彩铃和音频彩铃混合提示，并支持将彩铃类型返回给前端。基本规则如下：
2. 运营平台联想词库类型增加音频和视频彩铃类型，搜索平台提示时支持优先匹配联想词库的内容。无维度联想提示不匹配这两种类型的提示词。
3. 相同匹配度时，视频彩铃优先于音频彩铃；
4. 若音频彩铃的提示词包含于任何一个视频彩铃标题中，则给该提示词返回视频彩铃的类型。如：搜索“安家”，有个“安家”的音频彩铃精确匹配，优先级较高，有个视频彩铃叫“安家视频彩铃”。由于“安家”包含于“安家视频彩铃”，所以返回提示词“安家”时给与视频彩铃的类型。原因在于：后续前端会根据后台返回的类型自动跳转，如果不给定“视频彩铃”的类型，搜索结果中会自动切换到“音频彩铃”tab下。
5. 视频彩铃提示词的裁剪和排序规则复用5G全面屏。

### 1.2 搜索结果优化

【需求描述】

1. 视频彩铃支持无结果推荐
2. 当视频彩铃搜索无结果时，支持返回无结果推荐数据，由客户端控制返回，默认不返回。
3. 运营管理平台增加视频彩铃无结果推荐配置。

1. 视频彩铃取消歌手别名分词模糊匹配
2. 对于歌手别名，仅支持精确匹配；
3. 当搜索词包含字母时，要求搜索词分词个数大于1才支持匹配歌手别名。

badcase：



## 3.支持指定最佳展示位类型

【需求描述】

当前：最佳展示位只配置内容，不区分是A区的还是B区的，由搜索平台根据内容类型来判断该内容所属区域；

改为：在运营平台配置最佳展示位内容时，支持指定该配置是A区的，还是B区的，默认不指定；当运营平台指定区域后，搜索平台直接使用该配置，否则，还是按照现网的逻辑判断。

## 4.试听飙升标识

【需求描述】

若某个歌曲日试听增量超过500，则给与试听飙升的标识。该标识暂定一周的有效期，超过一周，删除该标识。该标识返回给前端的要求如下：

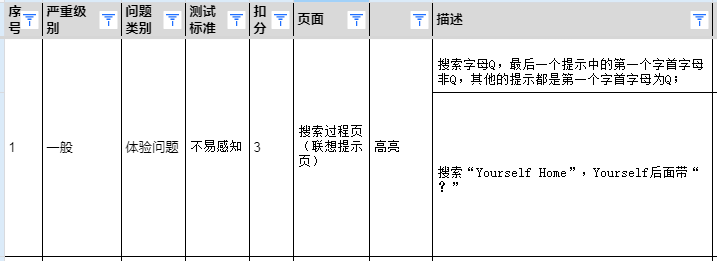
1. 该歌曲在非分词搜索结果列表；
2. 该歌曲有版权；
3. 该歌曲无点击率标识（超过XXX点击）和超多评论标识；

若上述三点均满足，则在原来的点击率标识字段中返回“试听飙升”的文案。

## 5.智能提示优化

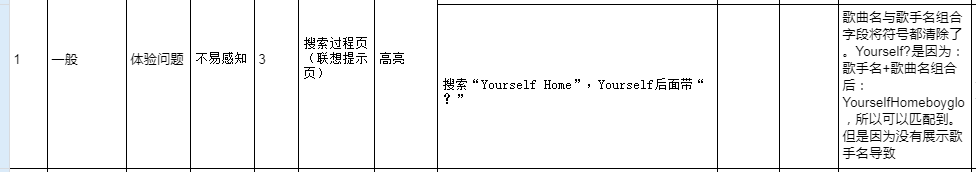
【需求描述】

1. 对于通过歌手名称召回的，需要提示出歌手名，包括“群星”、“various artists”等。组合提示不支持通过匹配“佚名”、“无”等召回。



1. ~~模糊匹配时，3.2.0版本要求了不返回组合结果。现在要求模糊匹配时~~不支持组合提示的场景（比如热度不满足要求、模糊搜索）不匹配songsinger和singersong。

Badcase：



1. 无维度智能提示增加过滤条件。对于歌手长度大于等于2的， **如果是以歌手名开头的**，在 歌曲+歌手 / 专辑 +歌手 组合的时候，如果有包含关系，取消组合提示，只保留长的那个。

Exp：



## 6.纠错的四个需求和一个问题定位

【需求描述】

1. 组合搜索支持别名全拼+汉字。精确匹配，不支持有错字、多字、漏字。

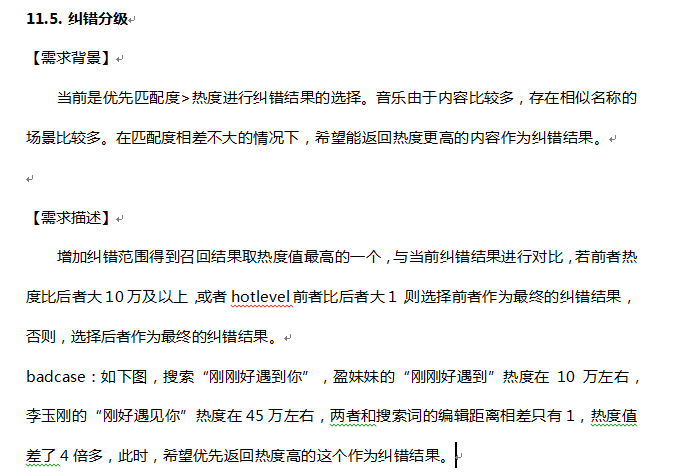
Badcase：“dengziqi 泡沫”、“dengziqi 摩天动物园”。

1. 组合搜索支持简拼+汉字。精确匹配，不支持有错字、多字、漏字。如果是别名的简拼，要求别名字数 >= 3 。

badcase： “zjl 晴天”

1. 纠错热度考虑hotlevel。实际hot = hotlevel \* 10万+hot；
2. 纠错分级优化。

背景：对于同一级的纠错类型，如果两个纠错结果热度相差10万以上，优先取热度高的而不是匹配度高的。详见v2.7.0 需求描述：



需求：数字转汉字优先于少字、多字、错字，即纠错优先返回数字转汉字结果。

badcase：

“199八”纠错为 1996，没有优先纠错为“1998”的问题。

【问题定位】

caankuyueguang linyoujia”期望搜到”残酷月光 林宥嘉“，实际无法搜到结果。

## 7.支持返回组合纠错词

【需求背景】

当前返回给前端的纠错词只有单一纠错的，当组合纠错结果有误时，用户无法点击按照原词搜索。

【需求描述】

对于组合有关联的纠错结果，若搜索词和纠错组合结果不一致，需要能将组合结果返回给前端。

Badcase：“zhoujielun 晴天”。

## 8.运营管理平台

### 1.1支持根据版权id查询歌曲/专辑/MV/视频彩铃

【需求描述】

数据查看菜单下增加版权id的输入框，支持根据版权id查询对应的歌曲/专辑/MV/视频彩铃。

### 1.2 联想词库增加有效期

【需求描述】.

数据查看——>联想词库。运营配置联想词时支持指定联想词的失效时间，超过失效时间，热度和运营热度都降为0，默认不设置失效时间。精确到天即可。

# V3.2.2

## 1.3D标识判断规则调整

【需求背景】

原有根据tonecontrol新增的几位来判断，由于音乐客户端对tonecontrol增加位数不适配，音乐临时调整了判断方式，后续等客户端适配后再改回根据tonecontrol来判断。

【需求描述】

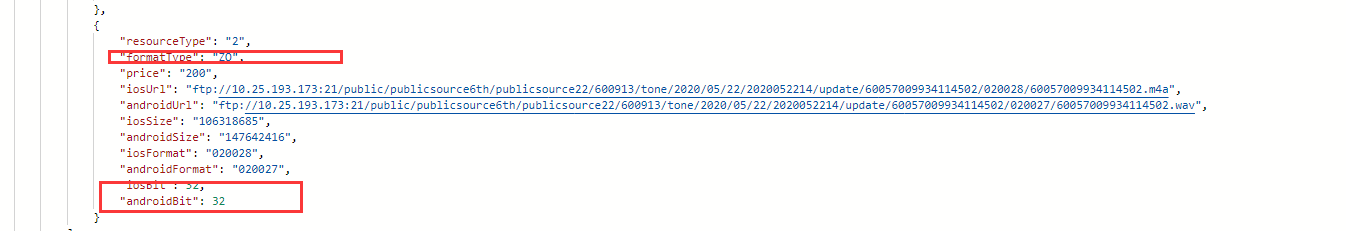
3D判断规则如下：先判断newRateFormats中是否有3D标识，然后根据搜索请求携带的客户端类型，分别判断IOS和Android的码率是否为空，不为空表示是3D的。

## 2.高bit至臻音质标识

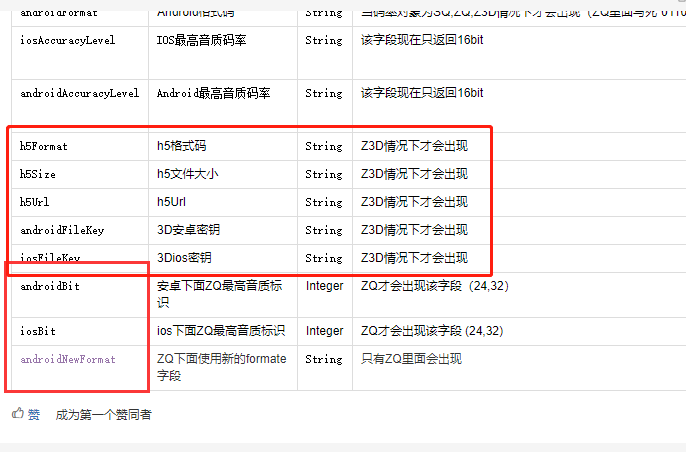
【需求描述】

高bit至臻音质标识判断规则如下：先判断newRateFormats中是否有至臻标识，然后根据搜索请求携带的客户端类型，分别判断IOS和Android的bit值。音乐后台只返回最高bit的值。

注意：原来24bit的判断规则也要改为上述方式。



注： 针对需求1、2，音乐接口返回报文中新增了若干字段，需要和视讯前端开发确认下有哪些需要透传：



## 3.热搜方案优化

【需求描述】

1. 运营平台支持增加热搜词，并支持设置该词的基础搜索量和类型（歌曲、歌手等）；
2. 对于人工添加的热搜词，采用基础热度+实际搜索量作为该词的搜索量，然后和自动的热搜词在一起排序并计算热度值。
3. 热词中返回的歌手数量可控，自动的+人工添加歌手总数不能超过该限制。

## 4.标识展示的若干规则变更

【需求描述】

1. 取消限制：当两首歌点击转化率相差10%以内时，两者的转化率都不展示的限制取消，只要满足要求都可以展示；
2. 增加限制：点击率标识要求只能展示在有版权歌曲上；
3. 增加限制：原有，无试听飙升标识展示个数限制；改为，限制最多展示两个且只在第一页展示。

## 5.运营平台支持厂牌数据管理

【需求描述】

内容管理——>上下架。支持厂牌数据的上下架。已下架内容在搜索结果中过滤。

# V3.2.3

## 1.增加SDK包名和签名校验

【需求背景】

当前的SDK包和后台交互时无包名和签名校验，导致SDK包有被滥用的风险。

【需求描述】

1. SDK获取客户端的包名和签名信息，并用和门户协商的加密方式，加密后携带给门户；
2. 运营管理平台。支持集成SDK的包名和签名管理。 在门户配置文件中增加报名和包签名的管理；
3. 搜索门户在收到搜索请求时，判断请求是否来自SDK，如果是，则对SDK传入的包名和签名解密后进行校验。校验通过，继续后面的流程；若校验不通过，则返回特定的错误码。

## 2.增加防刷量措施

【需求背景】

原有搜索接口缺乏必要的验证及限流措施，sign字段只针对appId和timestamp字段进⾏了签名，⼀直保持不变即可直接访问。

【需求描述】

将设备信息和用户信息加入sign进行签名，若相同设备/账号在短时间内大量访问搜索门户，则门户屏蔽此类请求，直接返回无结果。

1. 设备信息和账号信息由前端传入；部分设备信息需要用户授权才能获取到，客户端自己生成UUID作为设备唯一标识；
2. 时间限制和次数限制支持在配置中修改；
3. 运营平台支持白名单配置，在白名单中的账号不受刷量限制。

## 3.调用音乐接口时增加三点要求

【需求描述】

1. 调用详情接口是传参取消传入sid。原因：音乐侧反馈，sid会影响他们的缓存策略；
2. 调用详情接口的参数顺序根据参数名asc码值进行升序排序。原因：音乐侧会根据url来做缓存前置；
3. 调用热词接口的参数增加channel传参，参数由sdk调用接口时携带。

## 4.“搜索发现”模块热词个数限制

【需求描述】

搜索发现模块，热词限制最多展示6个，其中3个由运营配置，3个为个推。在调用音乐接口获取热词时，他们接口会返回热词类型。由门户限制个数。

注：如果该用户无个推，则展示6个运营热词。

# V3.3.0

## 1.配合前端改造或者涉及前端展示的

### 原唱标识

【需求描述】

根据歌曲标签中是否有“原创”标签、“原创单曲”标签的，返回“原唱”标签。规则如下：

1. 展示位置。与vip、音质等展示在同一级；
2. 展示优先级：下载标识 > VIP标识 > 音质标识（32bit > 24bit > SQ > HQ ） > 3D标识 > 原唱。

注意：“原创厂牌”这种不满足上述要求，无原唱标识。

### 最佳展示位支持功能搜索

【需求描述】

1. 运营管理平台最佳匹配位A区增加功能区配置，包括功能名称、功能描述、封面；
2. 搜索和门户将功能区类型返回给前端
3. 优先级：功能区优先级最低。



### 影视名露出

【需求背景】

当前部分歌曲的影视名字段存在多值，但是歌曲名/专辑名中仅展示一个影视名的场景。此时若通过另一个影视名召回歌曲，会让用户困扰。

【需求描述】

若搜索结果是通过影视名召回的，且影视名**不包含于**歌曲名（原名、别名）、专辑名（原名、别名）、歌曲描述中时，返回匹配的影视名字段，要求对影视名字段做拼接，拼接规则如下：

1. 判断原歌曲名中是否包含“主题曲”、“片尾曲”、“插曲”、“片头曲”“背景音乐”，如果有，则返回的影视名为：《影视名》+上述类型，如下：
   * 搜“亲爱的”，返回：隐形的翅膀 (电视剧《爱杀17》片尾曲) ===》 影视字段：《亲爱的》片尾曲
   * 搜“还珠格格”，返回：当 ===》 影视字段：《还珠格格》歌曲
2. 如果不包含上述5种类型，按照《影视名》+歌曲 作为拼接结果。
3. 最终返回给前端的为拼接后的结果。

### 歌手别名露出

【需求描述】

1. 涉及模块：最佳展示位、无维度搜索歌曲tab。
2. 若搜索结果是通过歌手别名召回，支持将歌手名拼接别名返回，返回规则如下：

* 要求别名不包含于原名（去符号匹配），否则不返回。如：周杰伦的别名“杰伦”，搜索“杰伦”时不拼接别名；搜索“she”时，不在歌手”S.H.E”上拼接别名“she”；
* 若存在多个别名，则返回包含该词的最长别名。如：taylor swift的别名包括“泰勒”和“泰勒斯威夫特”，当搜索“泰勒”时，拼接“泰勒斯威夫特”。
* 如果存在组合歌手，可不拼接别名。如搜索“jj”，某首歌为林俊杰和其他歌手合唱，则不拼接别名jj。
* 拼接规则为：原名（别名）。

### 标题简化问题处理

【需求描述】

1、待过滤内容字数判断指的是整体被删除部分的长度，如下，指的是红框内的长度，不仅是《》中内容的长度。



2、支持去符号匹配

目前看歌手的搜索结果未满足要求。



### 最佳展示位歌单播放量展示规范

【**需求描述**】

当歌单播放量超过1万时，播放量显示时以万为单位，小数点后保留1位小数。例如：当播放量为11280时，显示的播放量为1.1万。

### “搜索发现”模块热词个数限制

【需求描述】

搜索发现模块，热词限制最多展示6个，其中3个由运营配置，3个为个推。在调用音乐接口获取热词时，他们接口会返回热词类型。由门户限制个数。

注：如果该用户无个推，则展示6个运营热词。

### 非默认专辑名露出

【需求背景】

一首歌曲可能会被收录在多个专辑中，但是在搜索结果页中只会展示一个最优版权的专辑名称。

【需求描述】

如果歌曲是通过匹配了非默认专辑的专辑名召回的，将匹配的专辑名显示在歌曲描述中。（是展示在专辑右侧的歌曲描述，不是展示在下方的歌曲描述）

badcase如下：

搜索“明道 真爱”，排在前面的几个都是通过匹配专辑名“真爱”召回的，但是显示的专辑名不是“真爱”。



## 2.最佳匹配位优化

### 别名和前中缀匹配场景优化

【**需求背景**】

最佳匹配位，支持别名匹配，对于歌手、专辑、MV等，满足heatValue > 20万或者hotlevel > 0时，还支持前中缀匹配；

这就导致搜索“王俊凯”、“王源”等，歌手最佳匹配位出现组合“tfboys”，而不是对应的歌手之类的问题。

【**需求描述**】

1. 如果原名精确匹配结果为：1） hotlevel > 0或者heatValue > 20万或者isnew=1的专辑/MV；2） hotlevel > 0或者heatValue > 5万的歌手，则不管别名匹配召回结果热度如何，优先返回原名原名匹配结果；
2. 若歌曲搜索结果首位精确匹配歌曲/歌手/专辑，且歌曲hotlevel > 0或者heatValue > 20万或者isnew=1，则过滤最佳匹配位中前中缀匹配结果。
3. 若某一搜索词对应的最高歌曲转化率超过90%，则过滤最佳匹配位中前中缀匹配中无关的结果。

### 歌曲关联歌单取消分词匹配结果

【**需求描述**】

歌曲关联歌单的定时任务过滤分词匹配结果，此为临时方案，后续会根据用户点击情况对前中缀匹配的结果也增加限制。

## 3.歌手下关联歌曲搜索优化

### 综艺歌手的特殊处理

【需求背景】

当前音乐有一批虚拟歌手，是用综艺名称做歌手名。此类歌手只有与综艺专辑的关联关系，与歌曲名无关联关系，就会导致在SAM接口下，关联的歌曲为0，见下图：



【需求描述】

1. 运营管理平台支持给歌手设置是否综艺的标签；
2. 对于综艺歌手名，在SAM搜索时，若歌手下歌曲数量为0，支持将歌手下专辑内的歌曲填充到SAM歌曲搜索结果中。

注意：原有同名歌曲分散排序的要求仍然需要满足。

### 歌手SAM支持运营置顶干预（优先级低）

【需求描述】

对于运营管理平台。支持配置某歌手id下的歌曲榜单。支持配置置顶生效失效时间。

对于搜索平台，若存在歌手id置顶，则在SAM接口中优先选择置顶内容，再返回正常召回结果。

### 歌手下歌曲排序优化

【需求描述】

有版权组 > 无版权组。

组内按照：置顶 > isnew = 1 （isnew组最多一个） > hotlevel=6的（个数不限）> hotlevel\*10万 + heatvalue倒排。

注意：

原有的歌手下同名歌曲要分散展示的要求依然有效。

## 4.综合搜索优化

### 搜索截断

【需求背景】

当前音乐搜索精确匹配、前中缀匹配、分词模糊匹配等是在不同线程中执行的，当前是在等各个线程执行结束后将搜索结果按照对应的规则进行组合。这样会导致一个耗时较多的线程会影响整体的搜索耗时，且会召回一些不必要的结果。

【需求描述】

咪咕音乐APP歌曲搜索时（无维度的歌曲搜索），若精确匹配有版权搜索结果个数大于等于20，且第一位歌曲heatvalue>10万或者hotlevel>0或者isnew=1，则停止分词模糊匹配，直接返回精确和前中缀匹配结果。

增加场景：1、精确匹配小于20，继续模糊搜索（找一些歌曲结果少的精确匹配的歌曲）

2、首位歌曲hot<10万，hotlevel<0,isnew非1（可通过置顶实现）

3、首位歌曲hot<10万，hotlevel<0,isnew=1（可通过置顶实现）

4、首位歌曲hot<10万，hotlevel>0,isnew非1（可通过置顶实现）

5、首位歌曲hot>10万，hotlevel<0,isnew非1（可通过置顶实现）

6、繁简转换、汉字转换、转拼音同理

注：

1. 数字汉字转换、繁简转换、转拼音精确匹配都放在精确匹配逻辑中。
2. 需要加个开关，上线后有问题可以及时关闭该逻辑。

### 搜索分库分级

【需求背景】

当前音乐同步搜索数据较多，其中有很多数据一直无人试听或者长时间无人试听，此类数据被称为冷数据，反之被称为热数据。为了减少冷数据对搜索性能和效果的影响，需要对数据分库处理，冷热数据采用不同的召回和排序策略。暂定先只对无维度歌曲搜索执行分库操作。

【需求描述】

1. 歌曲入冷库规则

以入库时间为基准：

* heatValue小于1000且hotlevel<1；
* 入库1个月后播放总量为0或者近3个月播放量增量为0；
* 该歌曲对应的歌手非热门歌手（heatvalue>100）。

同时满足上述三个条件的歌曲添加到冷库中。

增加点：1、查看冷热库中数据是否满足各项条件

1. 歌手/专辑入冷库规则

* 歌手下关联歌曲数为0或者关联的歌曲都在冷库中；
* 专辑下关联的歌曲数为0或者关联的歌曲都在冷库中；



1. 分库搜索匹配规则
2. 对于热库内容。在当前的搜索规则上增加如下要求：

* 取消前缀匹配热度限制； 青春修（练手册）

增加点：修改歌曲热度后搜索该歌曲前缀，看是否显示在返回结果中（热库中）

* 支持模糊匹配/转拼音匹配热门歌词库的内容，若相似度>75%，支持按歌词库中的词搜索。 模糊匹配歌词的，要求歌曲 heatvalue >20万或者hotlevel>0才能召回，优先级等同于歌曲名模糊搜。
* 支持补词搜索：
* 搜索词存在于top1万热搜词中，查看该词的点击行为：若点击量最高的结果（可能没有对应的点击量数据）和搜索词不精确匹配（去括号、符号后的精确匹配），而是包含关系，且点击率大于50%，则搜索时将用户点击的结果的名称作为搜索词的补词结果，然后将补词作为容错词进行搜索；

增加点：

* 搜索词不存在与top1万热搜词中（去符号精确匹配），则从top1万热词中查看是否有包含该搜索词的热词，如果有，选择转化率超过70%的热搜词中搜索量最高的词作为补词结果，然后在前中缀搜索逻辑中将补词作为容错词搜索。（看一下转化率的表）
* 上述两种场景容错词匹配的结果，要求heatvalue>20万或者hotlevel>0或者isnew=1。

注：可提前通过离线任务计算出该搜索词对应的补词结果。

【badcase】

* 搜索“taylor”，未优先返回“taylor swift”相关的结果；
* 搜索“靓颖”，歌曲tab下，只返回了部分标题带靓颖的歌曲，基本未返回张靓颖的其他歌曲；
* 搜索 “凉”，“凉凉”未召回。

1. 对于冷库内容。在当前的搜索规则上减少如下逻辑：（可以修改热库改成冷库）

* 匹配项缩减：不搜索影视名、词曲作者、歌词、不支持转拼音匹配；
* 冷库中的歌手和专辑不支持分词匹配；
* 取消分词模糊匹配；
* 不召回无版权内容。

增加点：匹配项可以修改热库中歌曲项进行热库变冷库，然后进行搜索，

1. 冷库热库召回结果融合。(冷库精确的放在热库分词组前面)。

增加点：冷库中的词如果精确匹配的话需要放在热库中歌曲前面

1. 分级保底搜索逻辑（可暂时不进行设计用例）

搜索无结果时，支持保底搜索，在现有搜索规则下，增加如下规则：

* 支持数字汉字转换后前中缀匹配；（badcase：3次元搜索无结果）；
* 支持使用歌词纠错结果搜索；（搜索词词长度大于等于4）
* 取消前缀匹配要求heatvalue>20万的限制。

### 分组规则优化

1. 别名降级

【需求背景】

当前分组间有强制排序：热门歌手组 > 热门专辑组 > 热门歌曲组 > 非热门歌手组 > 非热门专辑组。由于没有考虑分组间的热度，且热门歌手的分组优先级靠后，会导致搜索summer、she时，一些热门歌曲排序靠后。（当前是通过置顶才能将对应歌曲排在前面）

【需求描述】

在使用纠错返回类型确认热门分组时，若同时有热门歌曲组和热门专辑/热门歌手组，且歌曲热度 > 30万或者hotlevel >0，则增加以下判断：

* 歌曲匹配的是原名；
* 歌手/专辑匹配的是别名；
* *~~若歌手匹配的是别名，且heatvalue+hotlevel\*10万 < 歌曲heatvalue+hotlevel\*10万+30万；或者若专辑匹配的是别名，且heatvalue <歌曲heatvalue+hotlevel\*10万+30万；~~*

若上述两点同时满足，则优先返回热门歌曲组1，要求组内歌曲热度 > 30万或者hotlevel >0，限制最多6个。再按照热门歌手组 > 热门专辑组 > 热门歌曲组2等排序。

badcase：搜索“summer”，现有直接优先返回热门歌手组内容，导致歌曲summer排序靠后。summer为吉克隽逸非媒体常用称号。

注：后续可能会增加别名等级的概念，避免一个热门内容的不常用别名对搜索结果造成影响。





1. 根据用户搜索点击行为确认分组优先级

【需求背景】（分两种情况，有别名和有点击率）

当前用户点击行为仅用于组内搜索排序调优，可能会由于分组排序不合理，导致部分点击量比较高的内容排序靠后。

【需求描述】造数据使歌曲有点击率并有歌手别名及歌曲精确匹配项

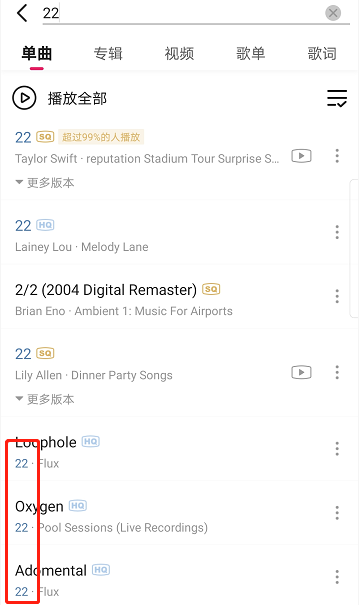
用户点击率超过60%的歌曲所在的分组为第一个热门分组。



1. 分组规则变更：

* 非热门歌手组增加限制条件：要求歌手heatvalue > 100或者hotlevel > 0（原有只限制了歌曲热度）。

badcase： 搜索“22”，歌手22热度为0，但是整体歌手组排序靠前。



* 取消非热门歌手/专辑分组的概念。统一作为非热门组进行排序。

### 抽屉排序优化

**【需求背景】**

部分场景下，抽屉中的排序不是很合理，会出现抽屉外的内容点击量较高，后续一直排序靠前，而抽屉内的一些更优质内容由于隐藏起来，点击量较低，会一直隐藏在抽屉中。

**【需求描述】**

如果抽屉内有其他有版权录音棚版歌曲的评论数远超过抽屉第一个歌曲的评论数（超过500），将该歌曲从抽屉中提取出来排在抽屉上方。

badcase：“清明上河图”，见下图



### 语义搜索功能扩展

【需求描述】

1. 支持词作者、曲作者 + 歌曲的组合搜索。如“周杰伦 淘汰”，能够搜索出陈奕迅唱的淘汰（是周杰伦作词作曲的）；
2. 支持歌手、词作者、曲作者组合搜索。如：“周杰伦 陈奕迅”，可以搜索出两人合作的“淘汰”；
3. 支持歌手+歌曲类型：“张学友 live”、“张学友 现场版”；
4. 支持歌手+音质：“周杰伦的3D 歌曲”，“张韶涵的24bit歌曲”；（优先级低）
5. 非歌曲tab支持语义搜索。包括专辑、MV、歌单、歌词tab的语义搜索。

* 对于专辑/MV：支持歌手+专辑/MV的组合搜索。如：“搜索周杰伦的魔杰座”
* 对于歌单：支持标签+标签组合。如：“推荐70年代粤语歌曲”；
* 对于歌词：支持歌手+歌曲组合，如：“周杰伦 简单爱”。

注：需要检查下“搜索周杰伦的魔杰座”，歌曲tab下怎么无结果。



### MV搜索优化

所有项增加无维度接口搜索结果

【需求描述】

1. mv搜索支持中后缀匹配，优先级暂定等同于分词匹配；
2. 模糊搜索的时候，如果是全英文的，支持根据空格分词。

badcase：

belong with me 搜不到 you belong with me

Give A Damn 搜不到 I Dont Wanna Give A Damn。

### 专辑搜索优化

【需求背景】

版权优于热度。即 在售数字专辑组 》热门有版权组 》 非热门有版权组 》 热门无版权组 》非热门无版权组。

有版权标准：专辑内有版权歌曲数>0；

热门组标准：hot大于20万或者hotlevel>0；

组内按照：isnew > publishtime > hot排序。



### 歌单搜索优化

【需求描述】

1. 歌单播放量参与排序

排序优先级：强版权热门歌单组 >强版权非热门歌单组 > 弱版权热门歌单组 > 弱版权非热门歌单组。组内按照播放量倒排。

强版权：有版权歌曲数大于歌单内歌曲总数的1/3；

热门歌单：歌单热度大于20万。



1. 歌单搜索支持匹配标签/标签别名

歌单搜索支持匹配标签和标签别名，标签名和标签别名优先级相同。别名数据参考如下运营平台内容：



badcase如下：来自于用户反馈。



1. 歌单tab支持繁简转换匹配&支持使用人工容错结果匹配

如题，要求歌单tab支持繁简转换匹配&使用人工容错结果匹配，优先级等同于原词。



### 歌词搜索优化

1. 支持繁简转换/数字转汉字/转拼音匹配

歌词tab下搜索支持繁简转换/数字转汉字补充搜索结果。

badcase：

* “吴克羣”搜索不出“吴克群”相关结果；
* “看我72变”部分“看我七十二变”内容搜索不出来；



1. 模糊搜索问题定位优化

我还是曾经那个少年

小朋友你是不是有很多问号

2、歌手、歌曲名精确匹配调优

歌词tab搜索 支持匹配歌手名和歌曲名，优先级同歌词。歌手名和歌曲名只支持原名和别名的精确匹配（原词精确或者纠错词精确匹配）。

注：歌词tab的话是要纠错类型是歌词的才拿纠错次去搜，如果不是，则只拿搜索原词去搜索结果

1. 有精确匹配结果后，若精确匹配个数少于10个，支持继续模糊搜索

badcase：“真的好想你 你会在哪里”



### 高亮的优化处理

【需求描述】

1. 对于通过精确匹配、前中缀匹配召回的结果，保持当前的逻辑（原词、纠错词、语义分析结果、过滤了单字的分词结果等）；
2. 对于分词模糊匹配的结果，只返回原词和分词的高亮词，对于单个字母，如果是名称中间的，返回高亮词时在字母右侧增加空格（单个字母的处理待定）。

### 支持全角半角转换

【需求描述】

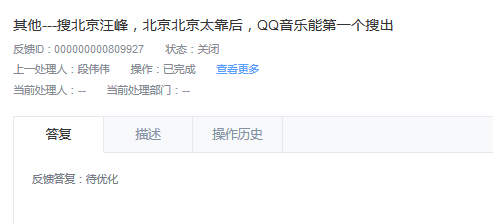
支持全角符号时可召回半角符号的内容，半角符号时也可以召回全角符号的内容。包括单曲、专辑、视频、歌单、歌词（修改歌词）。

Badcase：yourself？搜不到 yourself?

本版本只限于歌曲，其他类型不管

### 问题定位

1、搜索“北京汪峰”，排在前面的都是郝云的北京专辑里的。



## 5.音箱搜索优化

【需求背景】

当前音箱搜索和音乐app搜索使用不同的接口。音乐app的搜索排序逻辑已经优化了好几版，音箱搜索还是使用老的排序逻辑。

【需求描述】

将基础的排序规则整合成一个整体，音箱或者音乐APP搜索只做外围规则的限制，内部使用统一的排序逻辑。

## 6.智能提示优化

### 提示无结果时支持纠错

【需求描述】

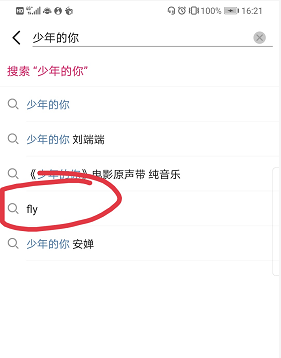
包括如下场景：

* 汉字+简拼。如：陈yx，提示出“陈奕迅”以及陈奕迅的组合提示；
* 全拼+简拼。如：chenyx，提示出“陈奕迅”以及陈奕迅的组合提示；
* 支持组合有关联。如“周杰伦的七里香”，支持提示组合“周杰伦七里香”；
* 支持多字/少字/错字。返回纠错结果即可，参考返回的高亮词。使用纠错词去匹配

### 针对带歌曲名（）的提示优化

【需求背景】

智能提示时会统一过滤标题中（）部分的内容，会导致通过匹配（）中内容召回提示时，会导致用户不知道提示词是怎么出来的。



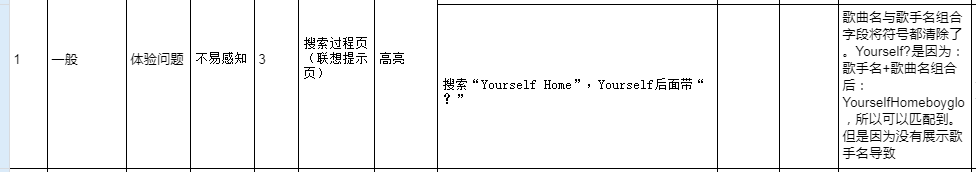
【需求描述】

1. 不支持匹配（）的内容召回提示词；
2. 智能提示时支持匹配影视名/综艺名，优先级与歌曲名相同。

### v3.2.1遗留问题处理

1. 不支持组合提示的场景（比如热度不满足要求、模糊搜索）不匹配songsinger和singersong。

Badcase：

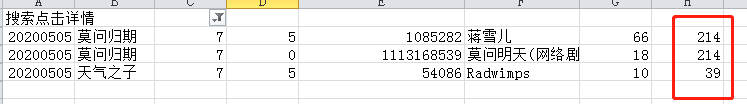


## 7.纠错优化

### 基于用户搜索点击行为的纠错优化

【需求背景】

下表为top 1万热搜词中，用户在纠错之后继续按原词搜索（type=7），并点击了原词搜索结果的数据。最右侧为“点击按原词搜索” 的搜索量，右数第二列为按原词搜索后的点击量。



【需求描述】

基于用户搜索点击行为的纠错调整。即：将搜索点击详情数据中，用户按原词搜索量超过10的搜索词，加入纠错白名单。

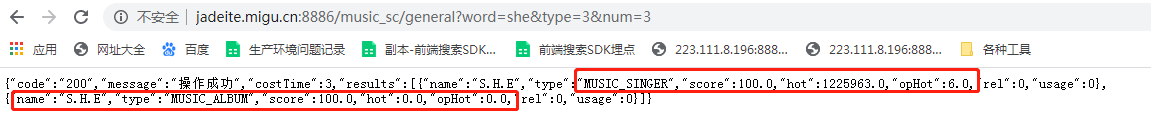
### 纠错结果过滤时清洗符号

【需求描述】

原有：当搜索词为字母时，若纠错结果返回不同类型的纠错结果不同，只保留和热度最高的纠错词相同的类型；

改为：判断是否相同时统一清洗符号后匹配。

badcase：搜索“she”,纠错返回的只有“S.H.E”的歌手和专辑，热门歌曲“she”未返回。



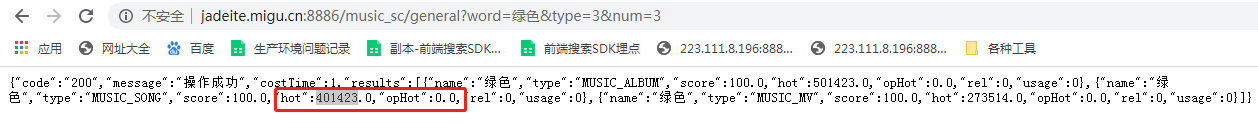
### 纠错返回歌曲选择

【需求描述】

当前：纠错返回的歌曲是按照heatvalue倒排选择歌曲；

改为：按照 hotlevel > heatvalue 的优先级倒排选择歌曲。

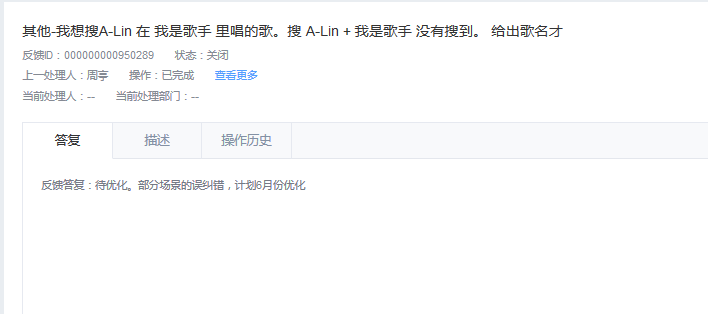
badcase：





### 问题定位处理

* NZND 未能优先匹配南征北战NZBZ；
* line 我是歌手，搜不到A-Lin在我是歌手里唱的歌；
* 组合无关联优先级低于 rel=0拼音汉字转换。如：“lin俊杰”，无法纠错为“林俊杰”的问题。



## 8.热度计算

### 热度计算支持匹配缩写&支持使用自有数据计算热度

【需求描述】

1. 在竞品上爬取元数据时，支持匹配缩写。 如 am（'m ）， is（'s），are（ 're）， not (n't )， is not（isn’t）、was not、are not（aren’t）、were not(won’t)、do not（don’t）、does not（doesn’t）、did not（didn’t）
2. 支持使用自有播放量、评论数等计算热度。具体规则待定。

### 歌曲热度计算支持匹配别名

【需求描述】

在爬取元数据计算歌曲热度时，支持别名匹配。

badcase：“赤子呓语一生梦 刺猬”，QQ音乐上歌手名是“刺猬乐队”，导致匹配失败，无法爬取到元数据计算热度。

### 增加爬取范围

【需求描述】

支持爬取抖音、QQ等榜单歌曲，将歌曲加入爬虫库。

## 9.数据处理

### 增加MV、视频彩铃的预屏蔽功能

【需求背景】

对于某些重点保障项目，咪咕音乐会提前分发部分内容，但是这部分内容不允许搜索露出。 当前仅有专辑和歌曲的预屏蔽，MV和视频彩铃提前分发的话，没法实时屏蔽。

【需求描述】

增加MV和视频彩铃的预屏蔽。支持在运营管理平台的预屏蔽词库中导入MV和视频彩铃的id。对于预屏蔽词库中的内容，不支持搜索露出。

### 32bit 优先于24bit

【需求背景】

搜索接口会根据数据处理计算出的音质得分，调整抽屉中歌曲的排序。

【需求描述】

32bit 音质得分高于24bit。

## 10.运营管理平台

【需求描述】

1. 支持根据上传者名称查询歌单

数据查看——>歌单——>增加一个输入框，可输入上传者名称，查询歌单信息。支持上传者和歌单名称组合查询。

1. 支持搜索干预项查询

支持输入某一搜索词，可以在页面显示该词是否经过人工运营，且支持显示运营类型，参考视讯运营平台功能。

注：即使只是给某个运营项添加别名A，则搜索A时，仍然可显示A对应的运营类型是什么。

1. 支持根据别名查询置顶配置

推荐管理——>搜索置顶子菜单，支持通过置顶别名查询置顶配置。

1. 增加热门歌词库管理

词库管理增加“热门歌词”子菜单，里面支持增删改查热门歌词，支持批量处理：

* 增：支持填入热门歌词即可（忽略大小写和符号）；
* 删：支持删除已配置的热门歌词；
* 改：支持修改歌词内容；
* 查：支持根据歌词内容查询歌词信息，包括歌词添加和修改时间，操作员等。

**待讨论点： 外文模糊搜索。**

# V3.3.1

## 1.实时热搜优化

【需求描述】

1. 自动热词增加过滤条件：（测试要在现网环境跟踪两天）；
2. 热词无搜索结果或者对应第一条结果为无版权内容，直接过滤；
3. 如果热词对应的第一条结果是通过分词匹配出来的，直接过滤；
4. 热词如果是运营平台配置的屏蔽词，直接过滤；
5. 实时热搜词中最多不能超过2个标签类搜索词，多余的过滤（个数支持在后台配置）；
6. 黑名单分级处理：

* 对于一级黑名单里的词：保持当前的逻辑，即：热搜词和黑名单中的词精确匹配时（去符号、去大小写），该词不能展示在实时热搜词里；
* 对于二级黑名单里的词：当热搜词包含同一个二级黑名单的词时，只保留最前面的热搜词。如设置“抖音”为二级黑名单，则热搜词“抖音热歌”、“抖音BGM”，实时热搜词里只能保留1个；
* 对于三级黑名单里的词：只要热搜词里包含三级黑名单的数据，直接过滤；如：设置“老歌”为包含匹配屏蔽类型，则“老歌经典”、“经典老歌”等都不能出现在实时热搜词里。
* 自动黑名单默认为一级黑名单，运营管理平台在黑名单设置时支持指定等级，默认是一级。

**注\*：同一个词只有一种等级，自动黑名单不能覆盖人工运营黑名单。**

1. 热搜榜更新机制调整

主要是解决两个问题：1）夜间分类热词较多异常；2）Top3霸榜问题优化。

* 凌晨1点~早上7点暂停热搜榜更新；
* 只有当前时段有效搜索量（门户统计的登录用户的去重搜索量（根据手机号去重））大于200的搜索词才能参与热搜榜计算；
* 除非是人工运营的热词，top3的热词不能存在长，

1. 优化热词类型统计不准确问题

【需求背景】

由于实时热搜有歌手名个数限制，该类型统计不准确会影响热词的展示。

【需求描述】

对于歌手、歌曲、标签等同名时，按照hotlevel\*10万+hot倒排，取热度最大的那个类型。支持匹配别名。

badcase：如“泰勒斯威夫特”被统计成歌曲类型。

## 2.合并非热门分组 – 330 遗留

【需求背景】

原有歌曲搜索排序有分组的概念：热门歌手 > 热门专辑…> 非热门歌手 > 非热门专辑 > 非热门歌曲…

【需求描述】

取消非热门歌手/专辑…分组的概念。统一作为非热门组进行排序，即：非热门歌手组不会强制优先非热门专辑组，以此类推。

## 3.搜索过滤“不展示”标签的数字专辑 – 音乐提

【需求描述】

1. 数据处理支持解析曲库同步的专辑数据的标签；
2. 搜索结果过滤标签中带有“不展示”的标签（搜索接口过滤 or 数据处理直接过滤都行）。

## 4.爱唱新版app支持数字专辑搜索 – 音乐提

【需求背景】

爱唱新版本客户端需要支持数字专辑，统一搜索需要做对应的改造。

【需求描述】

对于渠道为“miguyyTv”，且appname为“咪咕爱唱”，同时appver大于版本号X（具体多少和子公司确认）的产品，开放数字专辑搜索权限。

## 5.运营管理平台相关需求

【需求描述】

1. 实时热搜元数据展示

支持指定日期查询每个时段的top200的搜索词和其对应的Ht、Ht-1 ;

1. 实时热搜词结果统计

支持指定日期查询该天每个时段的热搜词以及热搜词对应的歌曲id和名称。默认展示查询当天的数据。

# V3.3.2

## 1.  【00750110294】搜索pieces时，无结果返回

# V3.4.0

## 1.搜索接口侧增加数字专辑和专辑下歌曲首发时间判断

【需求背景】

对于重点保障数字专辑，为了保证不因为DIH时间导致数字专辑延迟露出，会修改音乐同步的首发时间，提前计算。这就会导致，在无预屏蔽的情况下，数字专辑会提前露出。

【需求描述】

在搜索接口侧增加数字专辑和专辑下歌曲首发时间判断，即：数据处理侧可提前准备好数字专辑和对应的歌曲数据，但是不能修改原有的首发时间。搜索接口侧在判断数据处理已经准备好之后（iscomputer = 1），需要再增加判断是否到达首发时间，对于未达到首发时间的，不允许返回给前端。

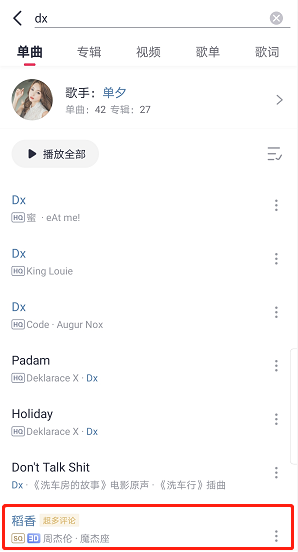
## 2.无维度歌曲搜索对纠错结果使用规则调整

【需求描述】

当前：当原词有精确搜索结果时，用纠错词搜索的结果就排序靠后，这就导致在使用简拼搜索时，部分场景下搜索效果不理想。

改为：对于搜索词为长度小于5的纯字母搜索，若纠错词hot > 20万或者hotlevel>0,则只有原词精确搜索结果hot > 10万 或者hotlevel > 1或者isnew =1，才优先展示原词搜索结果。否则，优先使用纠错结果（注意：此时需要在前端展示纠错提示）。

badcase：搜索dx，热门的稻香排序靠后；搜索sn，热门的少年排序靠后。



## 3.搜索结果中的pageIndex限制修改

【需求描述】

1. MV筛选接口/ SAM接口/厂牌专辑搜索接口，pageIndex限制最大为100；
2. SAM接口，支持根据查询到的歌手下歌曲数量限制pageIndex，即，若查询到歌手下歌曲数为800，每页最多展示10条，则调用SAM接口时pageIndex最大不能超过80，如果超过了，返回无结果。pageIndex不小于40，结果数total不大于1000；
3. 其他接口（不包括SAM）pageIndex限制最大为40；
4. 搜索结果数total最大为pagesize\*maxIndex，SAM接口上的结果数展示查询的结果数即可，不需要统计实际的搜索结果数。

## 4.搜索支持缩写匹配 -- 优先级低

【需求描述】

搜索时支持匹配缩写，如 am（'m ）， is（'s），are（ 're）， not (n't )， is not（isn’t）、was not、are not（aren’t）、were not(won’t)、do not（don’t）、does not（doesn’t）、did not（didn’t），使得用户在使用缩写搜索时可以搜索出非缩写内容，同理也支持用户使用非缩写的搜索词搜索出缩写内容。

注：

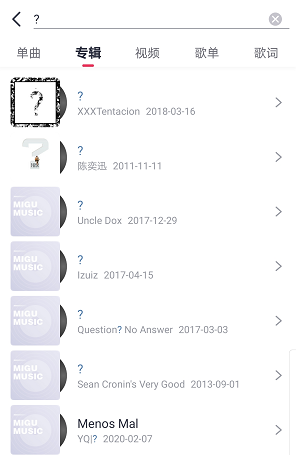
此版本仅在无维度歌曲搜索时支持。

## 5.专辑/MV搜索支持中英文符号转换 -- 优先级低

【需求描述】

上个版本在歌曲搜索时，增加了支持中英文符号转换的功能，本版本继续扩展，在专辑/MV搜索时也支持中英文符号转化（涉及专辑搜索、MV搜索、最佳展示位）。

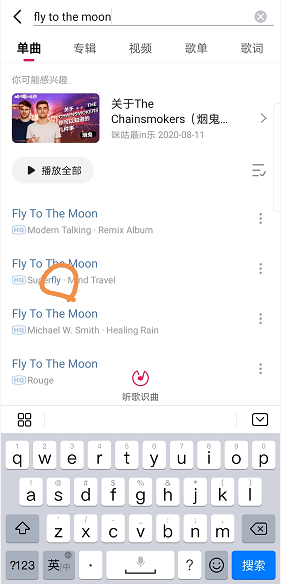
Badcase：搜索英文符号“?”，搜不出张韶涵的专辑“?”。



## 6.非分词匹配组高亮优化

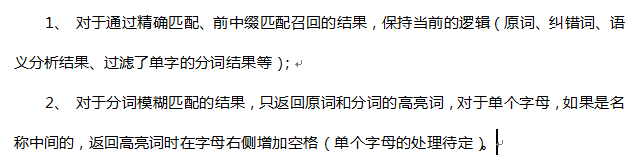
【需求背景】

当前通过前缀/精确匹配出来的结果，在部分场景下仍然高亮了分词，如下图所示：



【需求描述】

原有：



优化方案1：在原有1 过滤单字分词结果的逻辑中，增加要求：分词结果中多个连续字母作为单字过滤掉。

优化方案2：将原有1过滤单字分词结果的逻辑改为过滤所有分词。

## 7.歌词展示增加限制

【需求描述】

原有：若搜索词包含与歌手/专辑名中（去符号匹配），不返回歌词。

改为：若搜索词包含与歌手/专辑/歌曲名中（包括原名和别名，去符号匹配），不返回歌词。

badcase：16月6日。



## 8.热门歌曲组hot阈值修改

【需求背景】

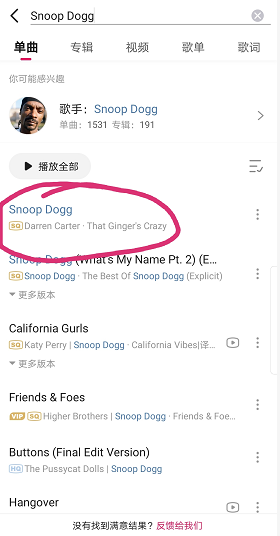
当前热门歌曲组的阈值较低，在无热门歌手/专辑组时或者热门歌手/专辑组6个歌曲之后，很容易有匹配歌曲名的非热门歌曲在搜索结果中排序靠前。

【需求描述】

热门歌曲组限制为：hot >0 或者 hotlevel >0 或者isnew = 1；

改为：hot >10万 或者 hotlevel >0 或者isnew = 1。

badcase：

## 9.爱唱增加伴唱搜索 – 音乐提

【需求背景】

咪咕爱唱伴唱搜索期望接入统一搜索能力。

【需求描述】

1. 伴唱数据同步

* 通过 6.16. IF2-16：门户接收通用素材接口，6.16.2.6. 咪咕爱唱对象：mgacMaterial同步爱唱素材；
* 通过6.22 门户接收编目信息的接口REST版同步编目信息；
* 通过6.25门户接收统一曲库分发咪咕爱唱产品接口 同步产品信息，以MATERIX444X，MATER开头的都是爱唱的版权。

1. 复用无维度接口支持伴唱搜索（无联想推荐）

* 支持根据歌曲名、歌手名（包括原名/别名）、歌手+歌曲组合搜索出爱唱伴唱内容，返回伴唱contentID和contentName；
* 复用咪咕音乐APP的排序逻辑；

注：

* 关于热度，根据伴唱中relatesong对应的热度，更新伴唱内容的热度；
* 爱唱搜索和咪咕音乐APP搜索共用运营平台同一个词库。

## 10.爱唱增加演唱会搜索 – 音乐提

【需求描述】

数据来源：MV素材中的演唱会类型，无码率限制；

搜索规则：复用现有MV搜索规则。

注：看是否可以复用5G专区1080P演唱会搜索接口（针对爱唱产品放开码率限制）。

## 11.支持12530接入视频彩铃搜索—音乐提

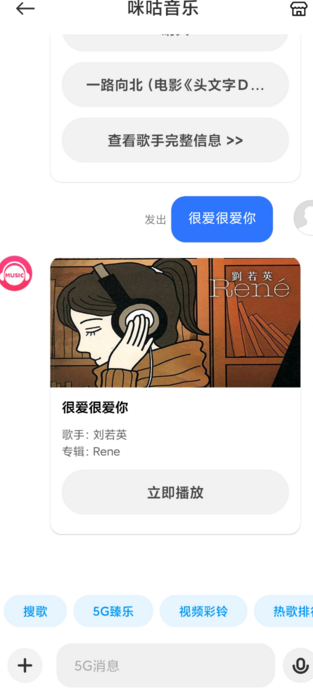
【需求背景】

12530准备接入统一搜索的视频彩铃搜索功能，复用当前的视频彩铃搜索接口，要提供新的appid用来区分不同的产品。当前视频彩铃搜索接口限制了appid，需要在代码中放开appid的限制。

## 5G消息歌手、歌曲搜索优化 – 音乐提

【需求背景】

5G消息里，允许用户输入歌曲名或歌手名，搜索结果包括两类：歌曲结果、歌手结果，两个结果不会同时展示，要么是歌手卡片，要么是歌曲卡片。该产品调用了无维度搜索接口，如果搜索结果中最佳匹配位有歌手，就展示歌手卡片，如果没有歌手结果，才返回歌曲卡片从左图可以看出，在部分场景下，上述逻辑会导致一些体验问题。

【需求描述】

针对5G消息这个产品，只有当用户搜索歌手意愿比较明确时，最佳展示位才返回歌手。避免人工运营（七里香，最佳展示位返回周杰伦）、补充（周杰伦 七里香，最佳展示返回周杰伦）对最佳展示位的影响。

方案：在返回搜索结果时，增加如下判断：1）纠错结果唯一&rel=0&歌手为热门歌手；2）否则，搜索词包含于歌曲列表第1位的歌手名/歌手别名中（支持去符号、去大小写匹配）。只要上述两个条件满足任意一条，即可支持返回歌手，否则，过滤最佳展示位的歌手。

## 13.数据处理

【需求描述】

1. 最佳展示位歌手下专辑数量统计过滤空专辑。
2. 数据处理过滤类型为“乐榜”的期刊：

【需求背景】

由于类型为“乐榜”的期刊不是原生的h5页面，搜索出来后，在前端点击进入详情页空白。

【需求描述】

数据处理过滤类型为“乐榜”的期刊，存量此类型的数据也需要清理。

# V3.5.0

## 1.普通专辑搜索增加上线时间限制

【需求背景】

歌曲有上线时间，数字专辑有首发时间，分别可以控制歌曲和数专不提前露出，但是目前普通专辑没法控制。因此，咪咕音乐曲库新增一个上线时间字段用于普通专辑上线时间控制。

【需求描述】

* 数据处理，增加解析专辑的上线时间字段；6.16.2.10.专辑对象：albumMaterial增加showTime专辑上线时间字段。
* 在搜索接口侧增加普通专辑的上线时间校验，如果该字段有值，且晚于搜索时的时间，则搜索结果中过滤该专辑。

注：

专辑和专辑下歌曲的上线时间是单独控制的，曲库那边进行专辑和专辑下歌曲上线时间的同步。

## 2.暂时屏蔽知识付费类数专和数专内MV

【需求背景】

咪咕音乐计划增加知识付费类数专（可以理解为听书类的专辑），但是目前原生客户端专辑详情页无法承载知识付费的专辑，以及播放详情页无处理此类知识付费视频（专辑下MV）。因此前期需要搜索进行屏蔽。

【需求描述】

1. 数据处理。屏蔽知识付费类数专。同普通数专一样，通过“不展示”标签屏蔽。
2. 屏蔽知识付费类数专下的MV。判断方式为：MV新增一字段（inDalbumFlag）用来区分是否为知识付费数专内MV，对于此类专辑，在搜索结果中不露出。（后续等客户端支持播放后再露出）；

注：

门户接收统一曲库分发MV接口REST版, 扩展字段增加isPaid是否付费、inDalbumFlag是否在知识付费数字专辑内。数据处理增加这两个字段的解析。（当前版本仅用到了inDalbumFlag）----这块的ispaid这个字段是不是不作考虑，只看mv的inDalbumFlag=1？

## 3. 适配咪咕音乐影视歌曲的歌曲名简化策略

【需求背景】

音乐计划对影视名歌曲名称做简化，即将原标题括号中的影视描述信息去除，统一更新到歌曲描述字段中。

【需求描述】

原有：如果是通过影视名召回的，会根据歌曲名中的影视信息拼接“影视名+片头曲/主题曲等”作为歌曲描述。

现在：由于改造后，歌曲标题里不再包含“片尾曲/主题曲”等信息，所以改为：当通过影视名召回歌曲时，优先返回歌曲描述中包含该影视名的描述。

注：

现在索引表里的影视名字段是通过提取歌曲描述《》中的内容，所以影视名肯定能找到对应的歌曲描述。如果因为返回了曲库删除了歌曲描述，就用《影视名》+歌曲作为歌曲描述返回。



# V3.6.0

## 1.定时任务更新榜单内容到置顶redis中

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750334688 | 在爬取咪咕音乐相关榜单后，将榜单内容添加到置顶redis中，置顶搜索词为榜单名称，歌曲id按照榜单中的顺序排列。后续每次更新时只替换置顶内容，原有的置顶搜索词或者其对应的别名继续保留。 |
| 需求规格：  无 | |

## 2.补充搜索时屏蔽“抽屉内容露出”的逻辑

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750334696 | 【需求背景】  原有将isnew=1的和评论数多的内容会从抽屉中单独提出来，以便于运营和排序自动矫正。上个版本在歌手名搜索逻辑过滤了上述处理，但是对歌手+歌曲/专辑的补充搜索逻辑中未过滤，导致搜索歌手+歌曲A时，在搜索结果列表里会有多个歌曲B未折叠。  badcase如下：  搜索 taylor swift my tears ricochet时，几首Back to december 未折叠到抽屉中。  IMG_256  【需求描述】  在执行歌手+歌曲/专辑的补充搜索逻辑中，过滤上述抽屉展开的逻辑，即：  屏蔽1：如果抽屉内有其他有版权录音棚版歌曲的评论数远超过抽屉第一个歌曲的评论数（超过500），将该歌曲从抽屉中提取出来排在抽屉上方；  屏蔽2：有新歌标识的，如果在抽屉中被折叠起来，需要将该歌曲从抽屉中提取出来。 |
| 需求规格：  无 | |

## 3.非精确有版权热门歌曲排序前置

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335210 | 当前： 强制精确匹配结果优先级高于前中缀和模糊优先级（精确组的最高热度大于0时），这就会导致很多弱匹配度的热门歌曲排序很靠后，如搜索：shake it，taylor swift的shake it off排在搜索结果五页之后。  改为： 优先排热门精确结果（要求歌曲hot>20万或者hotlevel>1或者isnew=1）；再将前中缀和模糊匹配结果中热度（hot+hotlevel\*10万）大于20万的歌曲取出6个，将热度减去20万后与精确组按照热度倒排。要求：模糊匹配的结果最多取两个且要求和搜索词只相差一个字（或者字母）。 |
| 需求规格：  无 | |

## 4.点击率标识展示增加限制

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335212 | 【需求背景】  当前点击率展示规则会导致如下问题：如下图，搜索“厚颜无耻”，展示点击率99%以上，实际上结果并不是用户需要的，此时的点击率标识会误导其他用户。  【需求描述】  为了避免透支用户信任，点击率标识增加如下限制条件：歌曲hot > 10万，或者hotlevel>0，或者评论数大于100。如果这些条件均不满足，则不展示点击率标识。  另：运营管理平台上点击率标识人工运营支持设置值为0。  IMG_256  IMG_257 |
| 需求规格：  无 | |

## 5.搜索无结果时支持纠错前中缀匹配

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335220 | 当搜索无结果时，支持将纠错前中缀匹配结果补充到搜索结果中（当前纠错词只做了精确匹配）。  badcase： 搜索“blgg”无结果。  IMG_256 |
| 需求规格：  无 | |

## 6.优化歌曲tab下歌词搜索逻辑

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335234 | 当前：歌曲tab下，即使只匹配了某句歌词中的一部分，也会将歌词召回结果放在精确组中，导致某些前中缀匹配歌曲、专辑等结果排序靠后；  改为：歌曲tab下，当通过歌词召回时，如果只匹配了某句歌词的一部分，将该召回结果放在前缀匹配组中；如果是精确匹配的才放入精确组中。 |
| 需求规格：  无 | |

## 7.优化冷库有原词匹配时的纠错提示

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335496 | 【需求背景】  当前纠错提示是否展示搜索原词未考虑冷库内容，导致误纠错时，用户无法通过点击原词搜召回冷库内容。  【需求描述】  当搜索结果中有冷库内容的时候，纠错提示都加上可点击原词搜索。（由于冷库不支持分词匹配，所以有冷库结果的时候点击按照原词搜索肯定有包含原词的结果）。 |
| 需求规格：  无 | |

## 8.支持搜索词局部替换

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335508 | 【需求背景】  1、当前简繁转换仅支持一对一，当一个简体有多个繁体时无法自动召回；  2、当某一综艺修改名称后， 不少用户还是按照原综艺名修改，综艺名第一季、综艺名第二季.... 如果完全让运营配置同义词或者别名会有露覆盖的问题。  为了在出现诸如以上的场景，增加如下需求。  【需求描述】  搜索词支持局部替换（类似于数字汉字转换）。  1、支持在运营管理平台配置原词和其替换词（支持多值）；  2、当搜索词中包含运营平台配置的原词时，支持同时用用户搜索原词和替换之后的词同时搜索，搜索原词优先级等同于替换词。  如：用户搜索词为“ABC”，运营平台配置了“B”可替换成“D”，那就同时拿ABC和ADC去搜索。 |
| 需求规格：  无 | |

## 9.联想提示若干优化

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335509 | 1. 当通过songsinger匹配时，直接提示songsinger，而不是先提示song、再拼接songsinger。   badcase：xhy，优先提示了邂逅；  850e8ac94e50b66364e03136d851f91   1. 在汉字转拼音匹配时要求召回结果第1位与搜索词相同。   badcase：搜索笑纳，提示出了“小奶狗”、“小奶瓶”等；  1603086756(1)  3、去除2次搜索中的分词匹配逻辑（与第1次的重复了），最多补充1个分词匹配结果。 |
| 需求规格：  无 | |

## 10.纠错词典过滤iscompute为0的内容

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335556 | 由于iscompute为0的内容在搜索结果中是过滤的，所以需要纠错词典过滤iscompute为0的内容，避免纠错后搜索无结果现象。  Badcase：  提示“井胧 想死却又不敢”是通过纠错返回的，实际并无这个可搜资源。 |
| 需求规格：  无 | |

## 11.语义搜索支持匹配别名和组合无关联补充

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335565 | 【需求描述】  1、语义识别出无关联实体后，支持根据“歌手+歌曲”、“歌曲+歌手”两种类型走补充逻辑。当前主要只考虑歌手+歌曲组合即可。以解决“江南  周杰伦”这种排在前面的全是周杰伦的歌曲的问题。（备注：当前语义识别出实体后直接按照热度排序的，是否要改成和普通搜索一样的排序逻辑还待定）  IMG_256  2、语义搜索支持匹配别名（先限制去符号大小写精确匹配），解决测试体验问题中的“查找邓紫棋的毒苹果”搜不到对应结果的问题； |
| 需求规格：  无 | |

## 12.智能补词增加限制

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335568 | 智能补词增加限制，要求补词结果的长度不能超过原词长度的3倍。（原搜索词长度暂不限制）。  badcase：  搜索zh，排在前面的很多都是周杰伦的结果。 |
| 需求规格：  无 | |

## 13.运营平台增加展示全网热度计算元数据

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335215 | 运营管理平台数据查看菜单增加“热度元数据”子菜单，该菜单页面有歌手/歌曲两个tab页：  1、每个tab下展示歌曲/歌手id、歌曲/歌手名称、hot、元数据（qq评论数、网易评论数、虾米播放量，没有就展示0）；  2、支持根据歌曲id/歌曲名 查找歌曲元数据；支持根据歌手id/歌手名查找歌手元数据。 |
| 需求规格：  无 | |

## 14.一些现网体验问题优化

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335617 | 【需求描述】  1、繁简转换中，“泪”的繁体由“淚”改为“涙”，避免“血汗泪”搜不出“血、汗、涙”；  2、最佳展示位逻辑：最佳展示专辑搜索结果中，如果有在售数字专辑，则优先返回数字专辑；无在售数字专辑情况时，再走满足热度条件优先返回原名精确匹配结果的逻辑（在v3.3.0）；  3、当用户点击按照原词搜索时，最佳展示位不再匹配简拼。badcase：  搜索“aq”，纠错为“阿悄”，最佳展示位展示“阿悄”，点击按照原词搜索后，最佳展示位简拼匹配了非热门的“獒犬”。  x31b0787e1b4074ec0b159896d0f3c67f5a64144249a3ac206d776ffb3e5fdf |
| 需求规格：  无 | |

## 15.纠错时别名降权

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335618 | 用户输入简拼，通过纠错返回汉字时，降低简拼匹配别名的权重，具体要求如下：  若简拼匹配的原名是 1） hotlevel > 0或者heatValue > 20万或者isnew=1的专辑/MV/歌曲；或者 2） hotlevel > 0或者heatValue > 5万的歌手，则不管简拼匹配的别名热度如何，优先对应的原名作为简拼对应的纠错结果。  badcase：  搜索“hl”，纠错结果返回庾澄庆的别名“哈林”，而不是热门歌曲“后来”。 |
| 需求规格：  无 | |

## 16.屏蔽词策略调整

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750335719 | 原有：当搜索词精确匹配（去符号去大小写）屏蔽词库中的内容时，不返回联想和搜索结果。  改为：  1、原有的精确匹配（去符号去大小写）改为增加统一转简体，使得配置简体/繁体后，对应的繁体/简体也能自动屏蔽；  2、屏蔽词典增加部分匹配类型，当搜索词包含此类屏蔽词时，不返回联想和搜索结果。 |
| 需求规格：  无 | |

# V3.6.1

## **1.纠错问题修复**

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
|  | 针对拼音搜索中字少、有模糊音及拼写有误产生的纠错问题进行修复。  Badcase：  搜索“kepule”没有纠错为“克卜勒”，而是纠错成无热度的“克鲁德”  *a70cc5dbab5ec9e5a724906874fcf03* |
| 需求规格：  无 | |

# V3.6.2

## 1.搜索接口过滤无全曲版权作品

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750337727 | 搜索接口过滤无全曲版权作品，涉及接口如下：  1、联想提示；  2、无维度搜索、歌曲搜索；  3、SAM接口。 |
| 需求规格：  无 | |

## 2.原唱标识同步

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750337697 | 咪咕音乐cms暂定12月12号上线原唱标识，为了避免后续处理存量数据工作量大，搜索计划同步升级数据处理逻辑，支持解析原唱标识，用于后期的原唱icon展示。 |
| 需求规格：  无 | |

## 3.支持搜索彩蛋功能

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750337696 | 【需求背景】  按照悟空部署的12月M-ONE推广计划，圣诞节期间（12月15日-12月25日）需要在客户端搜索栏上线输入口令展示掉落圣诞礼物的动效。  【需求描述】  当用户输入口令后，前端展示随机动效。  1、 运营管理平台   * 支持口令配置，即关键词配置，支持多值； * 支持配置口令的生效失效时间（精确到分）； * 支持配置同一设备动效可展示的次数，默认不限制次数，只要搜索关键词即可出动效； * 支持配置口令关联动效需要的图片地址（支持多值）以及动效出现的概率； * 支持配置图片掉落个数和速度。   IMG_256  2、 搜索接口/门户   * 需要在进入搜索启动页（实时热词所在页面）时将动效需要的图片地址返回给前端预加载； * 当搜索词命中任意口令时，根据运营平台配置的概率从图片地址中随机选择一个返回给前端。   3、 客户端   * 当后端返回图片地址后，前端可从顶部展示图片掉落的动画（路径随机），参见微信输入“生日快乐”后的动效； * 动效展示时长暂定为3s钟； * 在动效展示过程中不影响原有页面的操作； * 若后端设定该口令对应的动效在一个设备上只能展示一次，则当用户再次搜索时，不再展示该动效；否则，仍然支持展示。当一个动效未展示完，再次搜索时原有动效取消，重新展示动效。 |
| 需求规格：  无 | |

## 4.通过客户端和搜索sdk版本号减少刷量风险

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750337639 | 【需求背景】  sdk在3.2.0版本和搜索门户配合优化了防刷量措施，现在通过分析搜索日志，发现请求中会仿造sdk版本号绕过校验。  【需求描述】  门户增加配置，通过控制客户端版本号与sdk版本号的对应关系增加防刷力度。由于转化率计算现在已经限制从客户端v7.0.0（集成sdk v3.3.0）以上开始计算，所以当前增加限制为：若客户端版本号大于等于7.0.0，则限制sdk的版本号也要大于等于3.3.0（注意下ios和android sdk的版本号前缀不一样），否则认为该请求为刷量请求。  以下为异常请求：  IMG_256 |
| 需求规格：  无 | |

# V3.6.3

## 1.修复需求

sam接口当一页请求数等于歌手下歌曲总数时，返回结果为空

# V3.6.4

## 1.歌手搜索结果过滤无头像；门户增加歌手接口；

## 2.歌手下歌曲、视频数量统计规则调整

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750338681 | 1、独立音乐人歌曲数量计算  原有：对于独立音乐人，会将作词作曲的歌曲统计到歌手的歌曲数量下，但是在SAM中不展示作词作曲的歌曲，会导致数量和展示结果对不上；  改为：歌手下歌曲数量计算时，不考虑作词作曲数量。  2、视频数量计算  原有：统计了所有有版权视频数量；  改为：统计所有有版权“在线状态”视频数量。 |

# V3.7.0

## 1.视频彩铃搜索优化

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750337774 | 1、数据处理  从数据集市同步视频彩铃的播放量、点赞数、评论数、彩铃订阅量。数据集市每日将数据同步搜索的ftp上；  2、搜索排序  根据上述数据计算自有热度 = 点赞数\*100 + 采用歌曲热度计算公式计算热度（评论+播放量） --- 公式暂定如此，并将自有热度加入视频彩铃搜索排序中。  IMG_256  IMG_256 |
| 需求规格：  无 | |

## 2.热度计算增加校验

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750337779 | 【需求背景】  当前爬取第三方数据的时候存在较大的失败率，因此会因为前一次爬取失败导致本次计算的增量很大，影响热度计算的准确性。  【需求描述】  每次爬取的时候增加一个标志位用于表示本次爬取是否成功，建议当标志位为成功的时候才计算增量，计算热度。 |
| 需求规格：  无 | |

## 3.支持多维度自动置顶

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750338799 | 为了进一步实现运营自动化，增加支持多维度自动置顶功能。支持在运营管理平台配置置顶搜索词（以及搜索词的别名）、配置置顶维度（复用当前人工置顶的维度，包括歌手、歌曲等）、配置内容入库的条件、配置置顶的排序规则。如：  1、设置“独立音乐人”为【歌手】置顶，置顶内容选择条件是歌手的singerNormal为1，且歌手头像不为空，歌手的歌曲数或者专辑数不为0，排序规则是 粉丝数 > hotlevel > hot ，最大个数限制设置为100；  2、设置“全面屏彩铃”【视频彩铃】置顶，置顶内容选择条件是彩铃纵横比是16:9，排序规则是 hotlevel > publishtime；  3、设置咪咕汇的置顶为【视频彩铃】，置顶内容选择条件是彩铃标题里包含“2020咪咕汇”，排序规则是publishtime > hotlevel。  注：  一个搜索词可能会对应多个维度置顶，运营管理平台开发的时候可复用咪咕视频上自动置顶的页面样式。 |
| 需求规格：  无 | |
|  | |

## 4.歌手tab支持语义搜索&增加头像过滤

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750338990 | 1、歌手tab支持歌手语义搜索（仅用于歌手tab）。语义维度包括 性别（男、女、组合）、标签、别名，如：  -- 别名+歌手名，如：奶茶刘若英；  -- 性别+标签，如：香港流行男歌手；  -- 标签组合，如：香港知名唱作人（地域标签：香港、职业：唱作人）；  注：  单个标签和性别的，可以通过自动置顶填充来实现。  2、支持歌手+歌手组合。如林俊杰、周杰伦，陈奕迅薛之谦；  3、歌手tab搜索时过滤无头像的歌手；  4、排序 ：  -- 优先人工置顶/自动填充置顶；  -- 优先按照匹配度排序。各组内按照hotlevel\*10万+hot > 粉丝数 > 歌手id倒排； |
| 需求规格：  门户新增了歌手搜索接口，也需要测试 | |

## 5.非热门内容不支持不完整纠错

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750338991 | 当前：只要满足条件的内容，都支持不完整纠错（即：纠错结果只是名称的一部分）；  要求：改为只有当歌手满足热度要求（hot > 0或者hotlevel > 0 ）或者歌曲/专辑/视频/歌词/影视名等满足热度要求（hot > 10万或者后台level > 0）才支持不完整纠错，否则入纠错词典时只入原完整名称。 （现有设定一方面是为了减少不必要的索引，提高纠错速度，一方面是为了避免这部分数据影响热门内容的纠错效果。）  badcase：  搜索“起风啦”，预期是可以纠错为"起风了"，目前返回“起风啦”，就是因为有个非热门歌曲“又起风啦”。 |
| 需求规格：  无 | |

## 6.特定场景下歌曲名模糊搜索支持按编辑距离匹配

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750338992 | 【需求背景】  在上个版本“非精确有版权热门歌曲排序前置”的需求里，当模糊搜索时，经常因为分词的原因导致部分结果无法召回，更无法优先排序，因此增加如下需求，以优化模糊搜索时的效果，增加热门内容模糊召回的概率。  【需求描述】  当纠错结果返回的纠错词rel=0，且类型为歌曲/专辑/视频/歌词/影视名时热度小于10万，模糊搜索支持按照编辑距离匹配，具体要求如下：  1、汉字大于等于2个字，数字字母大于等于5个字时才支持按照编辑距离召回；  2、仅召回编辑距离为1的结果，最多取两个，且要求召回内容热度大于20万（hotlevel\*10万+hot）；  3、仅原词搜索，不包括简拼、全拼；  4、注意按照编辑距离召回的和普通分词召回的，需要去重处理。 |
| 需求规格：  无 | |

## 7.使用正则优化影视类搜索词的搜索效果

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750339050 | 【需求背景】  原有正则表达式是放在纠错模块，后来因为影响其他搜索词的效果，该逻辑被取消，但是取消后会影响影视类搜索词的搜索效果。  【需求描述】  影视类搜索词的正则表达处理逻辑放在搜索接口。  正则表达式包括：  XXX + 主题曲/片尾曲/背景音乐/片头曲/原声/插曲；  若搜索词满足上述正则表达式，则搜索接口支持用过滤了“主题曲/片尾曲/背景音乐/片头曲/原声”之后的词去匹配影视名（支持去符号、去大小写、繁简转换），优先展示原词精确匹配的，再展示清洗之后匹配影视名的结果（清洗之后匹配影视名的放在前中缀匹配分组里）。  【测试重点】：  1、关注影视歌曲的搜索效果；  2、用例需要多找一些非影视歌曲名但是满足正则表达的词进行测试，避免此需求影响其他搜索词效果。 |
| 需求规格：  无 | |

## 8.最佳展示位组合无关联场景限制

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750339506 | 最佳展示位对于组合无关联场景，只取歌手类型的。  badcase：  极品飞车8，最佳展示位展示了专辑“8”。 |
| 需求规格：  无 | |

## 9.若干现网体验问题优化

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750339414 | 1. 歌曲搜索：纠错前缀补充的时候，如果有召回的结果需要设置纠错提示。   Badcase：搜“nengmianchangjiang”，纠错前缀补充结果时未设置纠错提示。     1. 组合场景优化   原有：组合场景时，增加了前缀匹配，默认调用对与专辑同名的歌曲提前的处理方法，该方法将导致组合场景为歌曲+歌手时，把其他歌曲提前。  改为：查询时，boost里对songName\_qp或songNeatName\_qp:"全匹配"的增大得分；同时限制当组合场景有专辑纠错结果时再调用对与专辑同名的歌曲提前的处理方法。  Badcase：搜“m滨崎步”时，其他歌曲排序在歌曲m前面     1. 数据处理对歌名、专辑名、歌手名、mv名中的不可截断空格替换为普通空格。   Badcase：搜“ice paper”，概率出现最佳展示位没有歌手结果  4、人工运营的容错、意图数据不作为高亮词返回  Badcase：泰密容错为泰蜜，结果显示高亮  66ee4b7217b3f55b9f479113970c983   1. 去除计算得数的转换   Badcase：搜“二十四+七”，最佳展示位是313专辑   1. 别名原名拼接问题修复  * 别名包含在原名中的不需要拼接原名。 * 对原名和别名进行统一的格式处理，即去特殊符括号之类的字符之后比较。 * 对原名和别名去括号、中括号、大括号再比较是否一致。   注：歌曲搜索的场景同样处理  Badcase：搜索fx，最佳展示位歌手f(x)拼接了别名(fx)     1. 修复词曲作者召回结果和其他非热门组结果拼接按热度倒排的排序不佳问题。   Badcase：搜“周杰伦”，周杰伦为词曲作者的歌曲排序在周杰伦演唱的歌曲上面。     1. 修复继续搜索和简拼纠错热门内容的代码逻辑冲突问题   Badcase：dx    9、解决因置顶内容仍参与抽屉排序，评论数超过500的单独提前造成的hotlevel和点击意愿高的歌曲排序靠后的问题。  Badecase：夜空中最亮的星，hotlevel和点击意愿更高，但排序在第二位    10、无维度接口当是默认搜索请求(搜歌曲时)才查询返回跳转标识，其他tab下不再返回跳转的数据和flag。  11、歌曲搜索下的语义分支中的标签搜索增加标签同义词的判断  12、修复带符号的歌曲搜索结果排序不佳或搜不到的问题  Badcase：搜索“同根生 《泰雅梦游》：〈相聚歌〉、〈战舞〉”，精确匹配结果排序靠后  bff079a16a351e663a41e9c16f4d8fc |
| 需求规格：  无 | |

# V3.7.1

## 1.支持哼唱识曲接入统一搜索

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750341600 | 给哼唱识曲分配一套新的appkey和appid。  测试时可用音乐最新的测，通过哼歌识曲场景验证效果。 |
| 需求规格：  无 | |

## 2.票务搜索返回地址改为场馆名称

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750340774 | 为保持和音乐首页展示信息一致，票务搜索时，票务地址返回“场馆名称”。  IMG_256IMG_256 |
| 需求规格：  无 | |

## 3.歌曲名中带“纯音乐”的歌曲处理为伴奏版

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750341668 | 数据处理针对歌曲名中带 "纯音乐" 的处理为伴奏版。 |
| 需求规格：  无 | |

## 4.现网体验问题优化

|  |  |
| --- | --- |
| 需求标示 | 需求描述 |
| 00750128604  00750128606  00750128652 | 1. 修复歌曲搜索搜索接口没对未达到上线时间的普通专辑进行过滤的问题。   Badcase：《唐人街探案3》电影原声,未到上线时间被露出。  502c2a73f52c7265dccf1f5ba931a58   1. 解决音频彩铃维度搜索时，搜不到只有音频彩铃版权的歌曲的问题。   Badcase：白月光与朱砂痣，音频彩铃维度搜索时未搜到对应歌曲。   1. 修复歌单置顶，搜索结果不稳定问题   Badcase：搜索“CBA球员爱听的歌”，搜索结果来回变化、不稳定  IMG_256IMG_256 |
| 需求规格：  无 | |