**修订记录**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本号 | 修订日期 | 修改内容 | 备注 | 修订人 |
| USS\_music \_v3.10.0 | 5月18日 | 1. 最佳展示位优化 2. 增加搜索结果可解释性 3. 联想提示优化 4. 语义搜索优化 5. 搜索排序优化   --组内排序优化  --歌曲tab下专辑匹配的权重降低  --支持热门符号强精确匹配优先   1. 支持搜索结果自动化比对 2. 纠错的若干优化 3. 运营管理平台 4. 现网体验问题优化&搜索技术优化 |  | 段伟伟 |

# V3.8.0

# 配合前端改造

## 实时热搜模块改造

【需求描述】



【需求背景】

原有的「实时热搜」替换为多卡⽚展示，已定义4种卡片样式：搜索飙升、热搜歌曲、⼈⽓歌⼿、视频热播。搜索门户和运营管理平台需要配合sdk进行改造。

1. 热搜词由原来的点击后支持搜索和跳转h5两种方式，扩展为搜索结果页可跳转的方式都要支持，包括视频、视频彩铃等。当返回类型为视频时，需要增加返回视频封面、视频对应的歌手。



2、热搜词的选取规则待定。

3、卡⽚种类、数量、顺序前端不做限制，均由搜索后台控制。

## 搜索活动直达

【需求背景】

当前搜索侧经常会收到音乐活动热词配置的需求，这些活动对应有h5页面。但是当前搜索这些活动热词都是先进入搜索结果页，再从最佳展示位进入活动页面，增加了路径，且此类活动对应的歌曲结果列表大多不好。

【需求描述】 (对应 sdk\_v3.7.0)

支持活动搜索直达h5页面或路由跳转原生页面。即：

1、运营管理平台配置专题活动时，支持指定该活动是否搜索直达；

2、对于在运营平台配置了搜索直达的活动，在搜索活动名称时，搜索平台和门户只返回该活动给前端，不进行搜索。给sdk返回直达标识；

3、对于此类活动，sdk直接跳转h5页面或者根据路由跳转原生页面（如果同时存在，优先跳转原生）。

注意：

平台侧要注意版本控制，该逻辑仅适用于sdk v3.7.0及以上的版本。

# 最佳展示位优化

【需求描述】

1. 专辑露出增加限制（非组合关联场景）

背景：当前有很多单曲发行的专辑，当存在同名歌曲时，经常会有一些单曲专辑展示在最佳展示位上，造成数据冗余展示；

需求：对于非在售数专，要求只有当专辑有版权歌曲数大于0，且歌曲总数大于1时才支持在最佳展示位展示。（人工运营专辑不加此限制）。

1. 根据amber点击数据过滤。（待定）
2. 专区/活动类型扩展

只是扩展活动/专区所属分类，当前有下图所示几种分类。新增要求：1）运营管理平台支持随时写入新的分类；2）所属分类可以为空。



1. 最佳展示位过滤无版权的知识付费数专。

# 增加搜索结果可解释性（部分已完成）

【需求背景】

1. 歌手tab下支持展示歌手别名，展示规则同最佳展示位歌手别名展示；
2. 专辑tab下支持展示专辑别名，展示规则同最佳展示位歌手别名展示；
3. 最佳展示位支持展示专辑别名，展示规则同最佳展示位歌手别名展示；
4. 支持展示曲作者。当通过曲作者召回时，支持拼接曲作者提示，即：“作曲：XXX”，词曲作者同时命中时，返回词作者。

# 联想提示优化

【需求描述】

1. 联想提示简拼/全拼匹配优化

-- 歌曲、联想提示库支持简拼匹配。联想提示库中支持简拼精确与前缀，歌曲表里只支持简拼精确。简拼与全拼保持相同优先级；

-- 联想提示库中6个提示词的排序要求：对于hot >9000或者hotlevel>100的，原词精确 > 原词前缀 > 全拼/简拼精确 > 全拼+简拼前缀（当前原词和拼音的优先级是一致的）。badcase：搜索“she”，优先提示了“谁说我结不了婚”；搜索“f”,“faded”等热门内容未优先提示。

-- hotlevel <1的歌手不支持简拼前缀提示（联想提示库中的除外）。Badcase：搜索“dx”，优先提示了“单小源”；

1. 对于关联了id的联想词，如果该id已下架，该联想词不展示；（与8中运营平台的需求配合实现）；-- 运营管理平台来控制
2. 匹配歌手名组合提示词时，不做多歌手+歌曲/专辑的组合。（已上现网）

# 搜索排序优化

## 组内排序优化

【需求描述】

主要有两点变化：

* 用点击次数代替点击意愿；
* 评论数干预排序。

（点击率大于20%且点击次数大于50）或者（点击次数大于500的），按照如下优先级排序：

1. 点击次数；
2. hotLevel；
3. is\_new；
4. 评论数大于50的，热度降序排放；
5. 评论数小于50的，热度降序排放；

剩下的（不满足上述点击率和点击次数要求的）同样按照1~5的顺序排。

## 歌曲tab下专辑匹配的权重降低

【需求背景】

为了减少歌曲tab下专辑和歌曲名匹配的交替出现的现象，且尽可能保持优质内容优先展示的原则，增加如下需求。

【需求描述】

在非热门组中，在搜索排序时，通过专辑名匹配的要将hot-20万（热度不小于0）之后再与通过其他内容召回的结果进行排序。（原因：一般通过普通专辑名搜索歌曲的比其他的要少一些）。

## 支持热门符号强精确匹配优先

【需求描述】

当纠错返回的纠错结果带符号（非重音符号、非空格）且与搜索词精确匹配时，增加如下判断：

* 若带符号的为歌手，且hotlevel>0或者hot>10万；
* 若为其他类型，且hotlevel>0或者hot>20万。

满足上述任意一点，则在歌曲tab搜索结果召回时：

* 热门组召回时要求必须是带该符号精确匹配的；
* 增加该精确匹配的匹配度得分；
* 增加热门组个数，改为10个（现在是6个）。

在专辑tab搜索结果召回时：

* 优先展示符号精确匹配的。

在视频tab搜索结果召回时：

* 对于hot>20万或者hotlevel>0的，优先展示符号精确匹配的。

【badcase】

搜索“S.H.E”、“T.R.Y”等搜索词时，经常有不带符号的“she”“try”排序特别靠前，如下图所示:

【涉及场景】

先在无维度-歌曲、无维度-专辑、无维度-MV和歌曲、专辑、MV搜索接口试点。

# 支持搜索结果自动化比对

【需求描述】

          1、取出点击意愿表里的所有搜索词对应top3；

          2、同时取上述所有搜索词对应搜索结果的top1；

          3、对比下搜索结果top1是否在点击意愿top3里，如果不在则将搜索词和搜索结果展示出来；

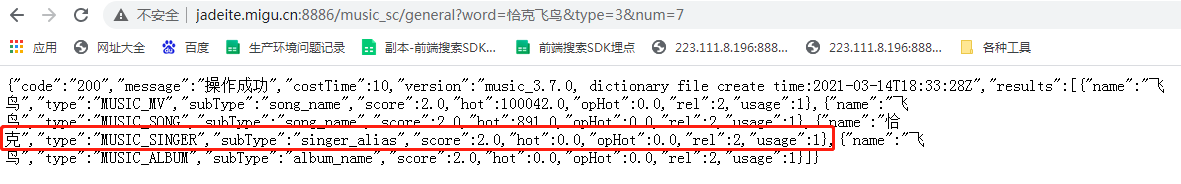
          4、每日对比一次，运营需每日检查。

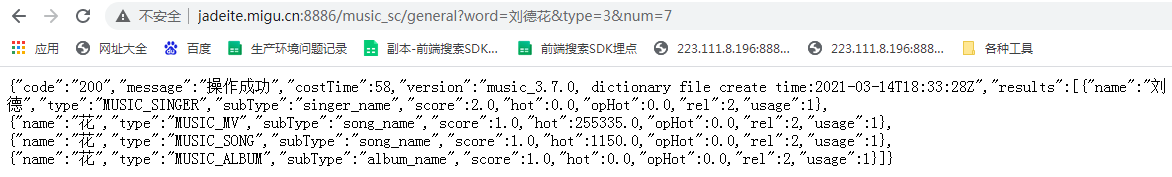
# ~~纠错的若干优化~~

~~【需求描述】~~

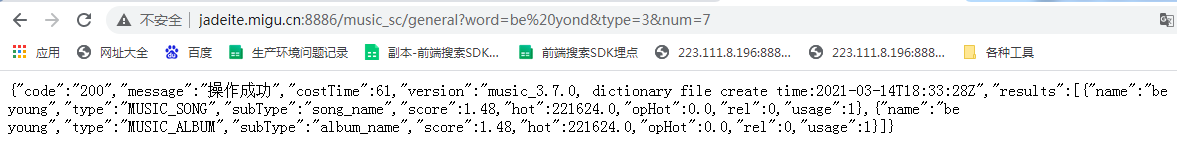
1. ~~组合无关联增加限制。当前会因纠错误返回的一些组合无关联结果导致搜索体验不好，因此组合无关联增加以下限制：若歌手（hot<100，且hotlevel<1），且歌曲/专辑/MV（hot<20万，且hotlevel<1）。纠错不返回无关联纠错结果。~~

~~badcase：搜索“恰克飞鸟”，返回如下组合无关联，期望结果是“恰克与飞鸟”；搜索“刘德花”，返回“刘德”和“花”；~~

~~~~

~~~~

1. ~~若某一搜索词包含于纠错词典某个词中，且词典对应的（歌手hotlevel>1或者hot>2万）；或者（歌曲/专辑/MV的hotlevel>1，或者hot > 30万），则不继续进行纠错。如果是精确匹配的，就返回对应的类型，如果是前中缀匹配的，纠错不返回结果。Badcase：搜索“张碧”纠错成“张磊”；搜索“chen”，纠错成“尘”；搜索“和乐器”返回“活跃期”，实际有个乐队别名叫“和乐器乐队”；~~
2. ~~带符号的纠错优化。存在空格场景时，如果结果是组合有关联结果，则优先返回组合有关联；否则单个符号错误与字/字母错误优先级等同。Badcase：搜索“be yond”，纠错返回“be young”。~~

~~~~

1. ~~乐队特殊处理。当歌手名/别名中满足“XXX乐队”时，将“XXX”加入白名单，避免误纠。~~

# 运营管理平台

【需求描述】

1. 支持电台数据管理

之前电台数据都是在redis里，未支持查看和运营。

* 数据查看支持根据电台名称、id搜索电台，页面展示包括但不限于电台id、名称，双击id可以查看详情；
* 电台支持删除（下架）；

1. 联想词库支持配置失效时间和关联id

* 失效时间精确到天即可，当天23:59或者次日0:00失效都行；
* 支持都某个联想词配置歌曲id，当该id失效后，该联想词也失效。主要是为了解决运营配置了歌词，但是该歌词对应的歌曲已经下架的场景，如：“早知惊鸿一场 何必情深一往”。

# 现网体验问题优化&搜索技术优化

【需求描述】

1. 热门组个数计算规则调整：

搜索结果支持抽屉时，热门组个数的判断改为按照抽屉外的个数来判断，避免因该组中存在多个抽屉或者一个大抽屉导致组内内容较少，badcase：she，热门歌手组里就3条数据。

1. 最佳展示位歌单展示限制

要求原最佳展示位没内容时才补充歌单展示。Badcase搜“winner”，最佳展示位没出现热门歌手winner，而是展示了包含歌曲winner的歌单；

1. 最佳展示位期刊按照资源发布时间过滤；
2. 最佳展示位支持专辑+多歌手搜索优化

为了避免有联想提示，没对应的结果，增加如下需求：

* ~~联想提示。匹配歌手时，只做单歌手+专辑的拼接，不做多歌手+专辑拼接；匹配歌曲/专辑时，最多只取前两位歌手与歌手/专辑组合songsinger/albumsinger。~~
* ~~搜索。如果专辑有多个歌手，增加一个专辑+顺序两个歌手的索引，这样当联想提示出现组合歌手时，可以保证对应的结果优先展示。只在最佳展示位和专辑搜索时支持匹配该字段。Badcase：搜索“《以家人之名》影视原声带 毛不易,马頔”，最佳展示位无该专辑。~~
* 搜索。组合搜索时（不是特指纠错的组合），最佳展示位展示的专辑歌手名中需要包含该歌手。Badcase：搜索“沈以诚《以家人之名》影视原声带”，最佳展示位上展示的专辑歌手名是“毛不易,马頔”。



1. ~~tag字段支持统一转小写去空格后匹配。Badcase：歌手tab下搜索“r&b歌手”，搜不到带R&B标签的歌手。~~
2. ~~搜索接口判断“原唱”标识时取消通过 “tag”判断的；~~
3. 歌词搜索过滤无版权冷库内容。Badcase：搜歌词“没有烟抽的日子”可以搜到，搜歌名“有烟抽的日子”搜不到。



1. 数据处理：

* ~~只保留音乐同步“原唱”标识的；~~
* ~~歌曲名括号（中英文括号）中包含Orchestral的，类型设置为伴奏版，和伴奏在热度减少、抽屉排序中的优先级相同；（目前看存量较少）；~~
* ~~清理歌手别名：女、男、女歌手、男歌手、热歌、热歌精选、热歌组合等别名信息不入索引表。~~