**修订记录**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本号 | 修订日期 | 修改内容 | 备注 | 修订人 |
| USS\_music \_v3.8.0 | 3月1日 | 1. 配合前端改造   --搜索彩蛋功能扩展  --最佳展示位视频化改造  --支持知识付费专辑展示  --搜索结果增加试听vip和付费标识  --SAM接口支持指定排序方式   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 | 预计3月31日上线 | 段伟伟 |

# V3.8.0

# 配合前端改造

## 搜索彩蛋功能扩展

【需求描述】

1. 支持彩蛋带链接

* 若元素配置了链接，在点击后，可跳转链接对应的h5页面内；如果无连接，增加点击动画；
* 支持配置每个活动单个用户点击到带链元素的次数。若用户跳转链接达到每日上限，则点击后不跳转h5；

\*注：

* 不同元素可能携带不同的h5链接，后端需要支持配置每个元素的链接，以及在该元素中出现链接的比例。比如搜索词A，对应元素B、C、D，B中出现链接的比例为1，则点击每个B元素都会出现链接，C中出现链接的比例为0.2，则10个C元素中，只有两个带链接，以此类推。
* 运营管理平台支持针对IOS和Android分别配置速度。门户根据请求来源返回对应的值。

## 最佳展示位视频化改造

【需求背景】

音乐2021年会进行视频化改造，增加优质视频露出，本次涉及最佳展示位的改造。

【需求描述】

1. 对于平台：

* 最佳展示位A区。运营未指定展区场景下，A区只展示新歌预约、歌手、功能，新歌预约>功能>歌手。若已经有一个展示在A区，则其他的都迁移到B区展示，否则默认展示在A区；
* B区支持同时返回多种类型，最多只能展示5条数据，除视频类型外，其他类型最多展示1条，只有在搜歌手时，才支持召回多条视频；
* 在B区中，歌手展示在首位，其他按照原先设置的优先级返回即可：

干预 >  歌手 >  在售数字专辑 > 普通专辑 = 最佳视频+10万热度 > 票务 > 演唱会 > 专区/活动 >  电台 > 期刊 = 歌单 > 彩铃（彩铃专区+音频彩铃+视频彩铃） 。

* MV支持根据歌手名搜索，限制MV hot >20万或者hotlevel > 0 或者 playnum大于1万。

1. 运营管理平台

* 支持控制某一内容在B区中的优先级；

注：

* 接口要注意适配老的sdk（版本进行配置）；
* 看下原先将歌手干预在B区的如何调整回A区。

## 支持知识付费专辑展示（最佳展示位+专辑tab）

【需求描述】

音乐综合搜索入口最佳匹配位+专辑Tab⽀支持展示知识付费专辑：

1. 对搜索平台：

* ⽀支持通过专辑名、标签（知识付费/知识）、专辑歌⼿手名、专辑内的音视频名称（只在专辑tab下）进⾏搜索；
* 该专辑不计入歌手专辑数量中，且不支持在歌手详情页搜索露出。
* 需要返回类型区别其他数专；
* 知识付费类专辑中的音频要过滤，不支持搜索露出；
* 只有sdk 3.6.0及以上的版本才支持此类专辑露出。

1. 对于门户

* 拼接知识付费专题跳转h5：

前缀+数专id，看是搜索平台还是门户给拼接下 <https://h5.nf.migu.cn/app/v4/n/knowledge-pay/index.html?id=600927015009000438>

* 专辑tab下当通过音视频召回专辑时，对应的提示语由“包含歌曲XXX”改为“包含XXX”。

## 搜索结果增加试听vip标识和付费标识（仅门户）

【需求背景】

音乐原对歌曲下载有权限控制（VIP和收费两种），即将对歌曲试听也增加权限控制；另外，歌曲最高音质和3D标识的判断方案也有变更。

【需求描述】

门户根据歌曲详情接口返回的showTag字段判断歌曲试听权限、最高音质、3D标识，该字段是一个字符串数组。具体方案如下：

该字段最多有5个值，其中a开头表示安卓识别，i开头表示ios 识别

* 第一位 用于判断当前歌曲的付费策略，有可能没有。等于VIP，则前端展示VIP角标；等于付费（具体取值看接口文档），则前端展示付费角标；
* 第二，三位， 判断当前歌曲的最高码率，（a24）安卓的最高码率为24bit，（isq）表示ios 的最高码率为sq；需要注意：hq不区分ios和Android，所以可能没有i和o的前缀，具体看接口定义；
* 第四，五位， 判断当前歌曲的是否有3D；



## SAM接口指定排序方式

【需求背景】

如下图所示，咪咕音乐歌手主页计划支持用户自己控制排序方式。



【需求描述】

SAM接口支持由产品侧传入排序方式，包括：最热、最新两种方式。不管指定哪种方式，有版权的永远在前面。

* 最新。优先按照歌曲发行时间（releaseDate）倒排，有发行时间排完了再按照creattime倒排；
* 最热。先选取50个歌曲置顶（如果没有50个就按实际数量选择）。优先取hotlevel=6的歌曲，然后取热歌榜中的歌曲，按照在榜单中的顺序选择，最后在热歌中（hotlevel>0或者hot>20万）按照前一日点击量倒排选取余下个数的歌曲。除置顶歌曲，余下的按照当前排序规则排序。更新频率，先设置为每日更新。
* 默认按照“最热”排序。

# 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。

1. 过滤（背景：当前有很多），
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。



# 增加搜索结果可解释性

【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”；

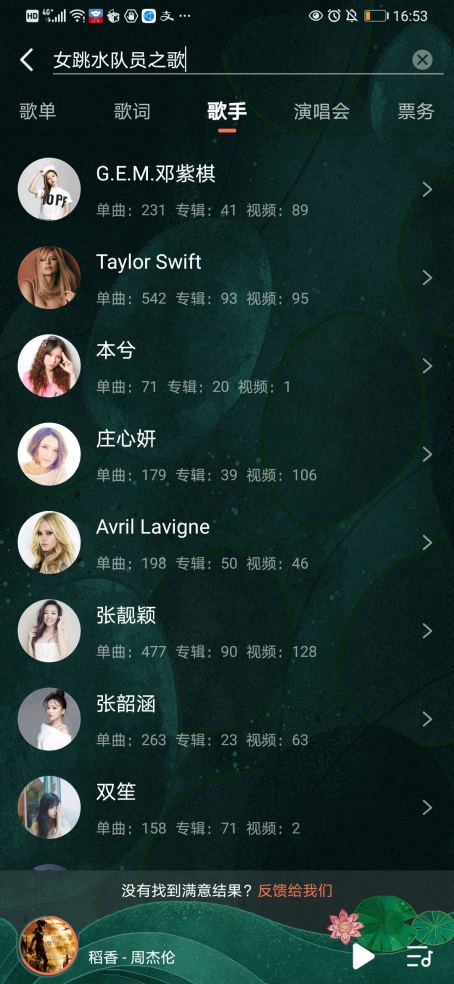
1. 对于关联了id的联想词，如果该id已下架，该联想词不展示；（与8中运营平台的需求配合实现）；
2. 匹配歌手名组合提示词时，不做多歌手+歌曲/专辑的组合。

# 语义搜索优化

【需求描述】

1、对于纯字母（包含数字）的搜索词，语义搜索只支持组合有关联场景，否则，只允许识别里面的歌手，且要求歌手的热度（hotlevel\*10万+hot）大于5万。badcase：搜索“you know will go get”，优先展示的全都是只匹配了will的结果。

2、歌手tab语义搜索，当识别出单一标签实体时，如果有其他非歌手的实体（如歌曲、专辑），则放弃本次语义识别结果。Badcase：



# 搜索排序优化

## 组内排序优化

【需求描述】

主要有两点变化：

* 用点击量代替点击意愿；
* 评论数干预排序。

点击率大于20%且点击次数大于50或者点击个数大于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】

1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>5万；
* 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；

【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；

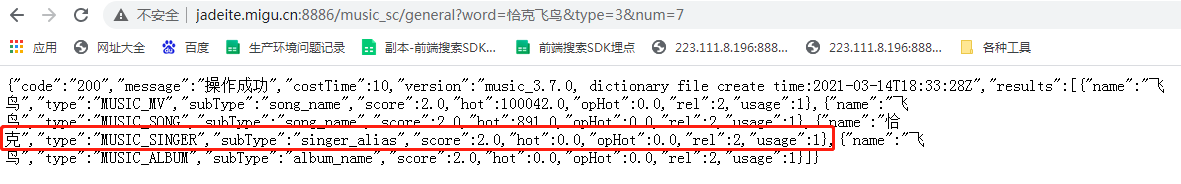
          4、每日对比一次，运营需每日检查。

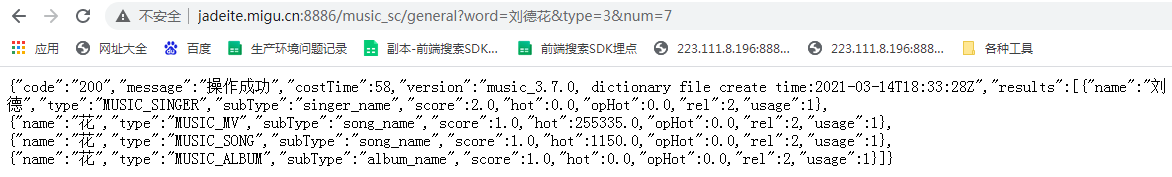
# 纠错的若干优化

【需求描述】

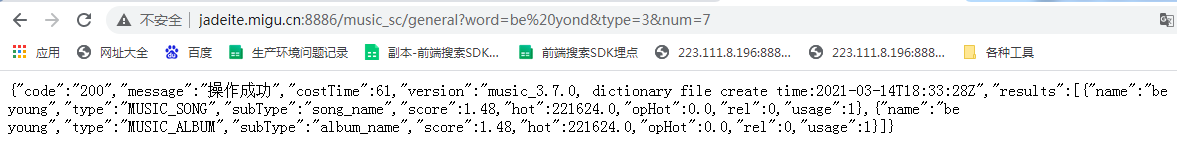
1. 组合无关联增加限制。当前会因纠错误返回的一些组合无关联结果导致搜索体验不好，因此组合无关联增加以下限制：若歌手（hot<100，且hotlevel<1），且歌曲/专辑/MV（hot<20万，且hotlevel<1）。纠错不返回无关联纠错结果。

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. 特殊拼音处理。当前“嗯”在纠错词典中对应的拼音为“ng”，导致搜索“ng”的时候会纠错成“嗯”，但是“嗯”在一般人的印象中都是读“en”，所以需要将纠错词典中的拼音特殊处理成“en”。
2. 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. 数据处理：

* 歌曲原唱标识去除通过“tag”判断的，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、男歌手等别名信息不入索引表。