

# Proposal for AI-Music system

## Abstract

我们试图用 PopMAG 模型，搭建便具有良好用户体验的 Web 端/微信端平台，为用户提供一个自动谱曲平台。

## 甜总的故事

- 甜总是一名音乐爱好者。

## 2020/10/8 晚

国庆假期结束了，甜总决定写一首歌纪念逝去的假期。坐在书桌前，甜总文思泉涌，很快写完了歌词。但是问题来了，甜总并不会作曲。

这时，甜总想起了最近听说的自动谱曲工具。他打开 AI Music 的网站，在网站上输入刚刚写完的歌词，点击“生成旋律”按钮。很快，网站上给出了一段根据歌词自动生成的旋律。甜总点击播放按钮，听了听生成的旋律，甚是满意，又点开这段旋律的 MIDI 数据，把 MIDI 数据保存了下来。

单有旋律可不行，还要有相应的伴奏才好听。甜总看着自己写的歌词，心情仿佛又回到了假期出游时的轻松愉悦，他觉得这首歌一定要是首欢快愉悦的歌才行。甜总在网站上选择歌曲情绪为“愉悦”，点击“生成伴奏”按钮，网站开始根据刚刚生成的旋律和甜总选择的情绪自动生成伴奏。过了一会，网站上生成了一段伴奏。甜总点击播放，房间里响起欢快的音乐，相比刚刚单一的旋律更加动听。甜总满意地保存了这段伴奏的 MIDI 数据，关上了电脑。

躺到床上，甜总哼起了刚刚的旋律，突然灵感乍现，对歌词有了新的想法。但是他已经懒得再起床打开电脑了，在微信里搜索自动谱曲，希望找到类似的小程序。甜总惊喜地发现，刚刚使用的网站竟然真的有相应的微信小程序！甜总躺在床上，把新的歌词输入到小程序中，根据新的歌词再次生成旋律。这一次，生成的旋律甜总更加满意，再次选择歌曲情绪为“愉悦”，为这段旋律生成伴奏。过了一会，相应的伴奏生成成功，甜总点击播放，跟着旋律轻轻哼唱，感叹自己的音乐才华。甜总将这段音乐的 MIDI 数据保存下来，将这段音乐分享到朋友圈。

甜总放下手机，哼唱着刚刚完成的歌曲，渐渐进入梦乡。梦里，甜总在自己的个人演唱会上，正唱着这首新写的歌.....

# AI-Music System 产品的分析

## 潜在用户

AI 生成的音乐虽然在艺术层面与人类演奏的作品仍有差距，但可以为以下用户提供帮助：

- 音乐表演艺术家，AI 生成的音乐可以为其提供灵感
- 歌唱家，AI 生成的音乐可以为其提供伴奏
- 音乐使用者，如影视、游戏厂商等，可以规避版权问题
- 个人音乐爱好者
- 声乐培训机构

## 产品形式

- Web 网站
  - 使用者可以访问本产品的网站或微信小程序获取服务
  - 优点：部署在本公司网站上，应用形式可控
  - 缺点：应用范围较小
  - 适用人群：所有用户
  - 微软利益：有更好的展示效果，可以直接得到用户反馈
- SaaS 服务
  - 使用者可以在其应用中调用本产品的 API，获取音乐生成服务
  - 优点：作为第三方服务，任何开发者均可使用
  - 缺点：若使用者较多，则维护难度较大
  - 适用人群：在线声乐培训机构、其它 Web 开发者
  - 微软利益：他人调用本产品 API 可能需要付费，有更大商业利益
- 产品设计
  - Web 网站
    - 业务简单，可能可以通过静态网站+后端调用模型推断的方法解决问题
    - 提供的功能比较单一，因此界面设计应突出主题，不必过于复杂
  - SaaS 服务
    - 需要对提供的 API 接口进行调研和设计
    - 如果用户比较多，后期维护可能存在问题
- 结论

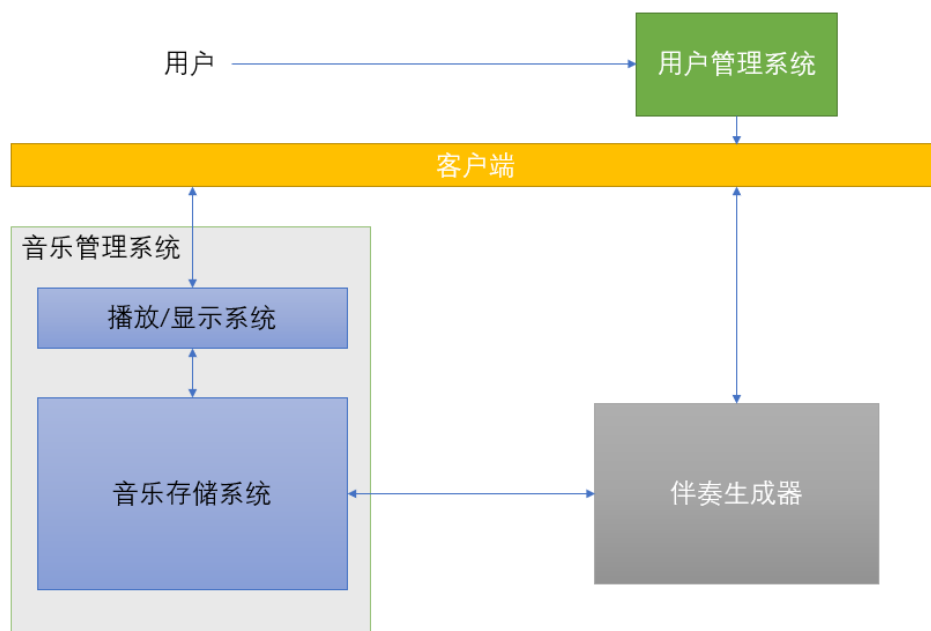
- Web 网站在本质上也是调用后台的 API 接口，因此两种产品形式本质上一致，但我们决定先从 Web 服务着手
- 本产品的 Web 界面逻辑实际上较为简单，最核心的是伴奏生成模型
- Web 网站需要设计浏览器端和微信小程序两种界面，它们可能有所不同，在技术选型时可能需要加以区分

在下文中，我们详细介绍本系统最核心的部分——伴奏生成系统 PopMAG。

## 伴奏生成平台 PopMAG Platform

该部分讲解我们的软件，伴奏生成平台。

软件的大致结构如下图所示：



- 系统功能
  - 输入：歌词
  - 输出：旋律加伴奏
  - 根据用户输入的歌词，生成相应的旋律加伴奏。用户可调节生成的伴奏和旋律的情感。
- 可用资源
  - 模型

- MSRA 提出的 PopMAG 模型(Ren, Yi, et al. "PopMAG: Pop Music Accompaniment Generation." *arXiv preprint arXiv:2008.07703* (2020).), 该模型可生成音乐伴奏。
  - 现有一些关于音乐生成的情感的问题的一些工作, 但是这一方面到具体使用还需要进一步的探索。
- 设备资源
  - 已有一台可用于部署深度学习模型的服务器
  - Azure 平台上的计算资源
- 宣传资源
  - 公司的官方宣传路径可协商用于宣传产品
  - 产品未来可在公司相关的 app 平台上架
- MSRA 科研资源
- 微软亚洲研究院的实习生可作为产品的第一批试用人员
- 产品形式
  - 在微信 app 中以小程序的形式上线
  - 作成网页版本上线
- 产品 Draft Spec
  - 用户管理系统 - User Management
    - 微信账号/微博账号登录
    - 不登入, 以游客形式使用产品
  - 音乐管理系统 - Music Management
    - 音乐存储系统
      - 以用户为单位进行管理。对每个用户分配一个单独的区域存储其相关信息
      - 在每个用户的区域内, 以歌曲为单位进行存储。将歌词, 生成的伴奏旋律等以一个结构存储在一起
      - 能够与其他系统进行交互, 允许特定的系统在这里开辟账号, 存储, 读取信息。
    - 播放/显示系统
      - 能够从音乐存储系统中读取旋律, 歌词等信息
      - 能对旋律, 歌词进行协同的播放和显示
  - 伴奏生成器 - PopMAG model

- 根据所给的歌词生成相应的旋律/伴奏
- 该部分需要 GPU 的支持
- 客户端 - client side
  - 微信小程序：能够在手机端，使用微信快捷使用，符合现在年轻对手机端的需求
  - 网页版本：适合于在电脑端使用系统的情况。无需安装，在任何地点都可以用浏览器打开
- 分享/合作 - Share/Cooperation
  - 用户可以将合成的结果分享到微信，qq 这些软件上，或该系统的其他用户
  - 可以多个用户在系统内协助完成一首歌曲的生成