

i-speak 口吃辅助治疗平台

作品设计说明书



i-speak
i-speak

让开口，不再是一种恐惧

目录

- 一、平台介绍..... 4
 - 1.1 平台概览与设计理念..... 4
 - 1.2 国内口吃患者及行业现状..... 5
 - 1.3 竞品分析调研结果..... 6
 - 1.4 平台的亮点和创新之处..... 7
 - 1.4.1 概念创新..... 7
 - 1.4.2 市场创新..... 8
 - 1.4.3 技术创新..... 9
- 二、特色功能概览及界面交互设计..... 9
 - 2.1 基础功能简介..... 9
 - 2.2 针对口吃语句的科学评分功能..... 11
 - 2.3 包含丰富语料模块的针对训练功能..... 13
 - 2.4 基于历史记录与音节解析的分析推荐功能..... 14
 - 2.5 论坛交友与经验分享的社交功能..... 15
- 三、核心技术路线..... 17
 - 2.1 基于定时采样和词级别分析的语音识别解析方案..... 17
 - 2.2 基于改进后的距离编辑算法的评价与分析数学模型..... 18
 - 2.3 基于内容推荐算法的语料库选择模型..... 21
 - 2.3.1 基于模块主题的推荐算法..... 21
 - 2.3.2 基于受阻音节分析的语料推荐算法..... 22

2.4 技术路线可行性分析.....	24
四、商业推广计划.....	25
4.1 市场调研及结果分析.....	25
4.2 目标服务群体分析.....	26
4.3 市场推广预测.....	28
4.4 SWOT 分析.....	30
4.4.1 Strength: 优势.....	30
4.4.2 Weakness: 劣势.....	31
4.4.3 Opportunity: 机遇.....	31
4.4.4 Threat: 威胁.....	32
五、结语：未来遐思.....	33

一、平台介绍

1.1 平台概览与设计理念

(1) 创意来源

在我们的前期实践调查中，我们注意到中国具有着庞大的口吃患者群体，很多口吃者会因为口齿不清晰问题而遭受各种不公正的待遇，这一点在幼年口吃患者的童年表现得尤为明显，同龄人的嘲笑甚至谩骂对他们的人生产生了持续且可观的消极影响。更引人深思的是，大多数口吃患者因为没有治疗渠道的信息，也没有合理的方法改善自己的口吃，口吃这一阴影往往会其伴随终生。

与此同时，一些口吃医疗辅助治疗机构通过专业医师的发音矫正与跟读训练，可以改善和治疗患者的口吃情况。很多口吃治疗理论在相关论文中也多有表述。我们能否把这种“医师指导下的发音矫正与跟读训练”在便携式实体上实现与应用，是我们开发这款口吃辅助治疗平台的灵感来源。

(2) 平台简述

本文档描述的 iSpeak 口吃辅助治疗平台，针对中国口吃患者群体，通过基于任务训练机制的针对性发音练习达到对口吃的治疗与缓解效果。平台通过建立口吃词句的评价与分析数学模型，使用语音识别与分析技术和基于内容的语料库偏好推荐算法，开发了流畅发音训练与评分功能、阻塞音节分析功能、针对训练语料库推荐功能、社区经验分享与交友功能等特色功能，帮助广大口吃患者群体改掉口吃习惯，更好地融入社会。

(3) 平台载体

项目现使用 Java/SSM 框架、Swift/Sqlite 框架等技术开发，应用主体目前已经实现跨平台（Android 平台和 iOS 平台）。为了更接近真人辅助治疗的效

果，我们在开发的智能互动机器人上也搭载了最新的 iSpeak 系统，通过扩展平台载体，来进一步提高平台实用性和平台治疗效果。

1.2 国内口吃患者及行业现状

我国口吃患者群体规模庞大，患有口吃语言障碍人数已达 1300 万人，大约占总人数 1%，同时我国每年口吃者的增长率率达到了 0.08%。

在中国，目前口吃都被归到教育的语言康复领域。这也是我们判断中国口吃行业，将来会归到教育的语言康复领域的一个非常重要原因。

在台湾和香港，目前已有很多大学在研究这个领域。如台湾大学语言康复学院，香港大学语言康复学院，香港中文大学语言康复学院，他们都对口吃有系统的理论研究，也产生了很多理论成果。

我国在 2003 年通过了语言康复立法，并成立了语言康复协会。在 2008 年左右，通过了语言康复师认证。中国的语言康复理论水平和国外比，还有一定的距离。香港在 2008 年才成立了语言康复立法，2016 年 12 月才成立语言康复协会。在口吃科研方面，几乎也没有什么理论著作。

在中国大陆，这几年已经也有很多大学及教授在研究口吃，如北京语言大学语言康复学院高立群教授，上海华东师范大学黄昭鸣教授，北京师范大学彭聃龄教授，清华大学基因研究实验室孙义民博士，首都医科大学康复学院李胜利教授和昆明医科大学熬丽娟教授。这些大学教授都对口吃有研究，也产出了很多理论和著作。由于一些原因，目前这些理论只在康复领域圈子内流传，圈子外的几乎看不到。根据目前掌握的调研资料看，主流医疗行业目前还没有把过多的注意力投放到这个群体上。

1.3 竞品分析调研结果

目前，口吃行业一直是一个处于社会关注度不高的行业。但仍有不少的口吃治疗师，口吃康复机构等存在。但国内并无官方权威的机构或者组织在引领整个口吃行业的导向与发展。许多由口吃者，或曾是口吃者的社会群体自发性的组织是这个行业的核心。但其中大多都是非盈利组织，更有与志愿者，爱心协会等合作的公益组织。

由于没有官方界定的行业规则 and 标准，导致口吃行业的门槛十分低，基本上没有什么明确的门槛的。因此造成行业的总体专业水平较低，许多诈骗、虚假、伪造的口吃治疗师或者口吃康复组织等机构在行业内活跃却不被大力的打压与制裁。这导致了行业内部混乱与难以健康有序的发展。

在现在，智能设备广泛进入大众生活的时代。几乎人人都有智能手机，而在国内却没有一款针对口吃群体的测评、治疗，社区一体化的移动应用程序。而本团队正是从口吃者受社会关注力度较小这一角度出发，研制一款结合时代背景与技术的智能系统。这对于中国国内目前的口吃行业而言，无疑将起到积极推进的作用。也将提高整个口吃者群体的社会关注力度。

我国已经进入老龄化社会，未来中国需要言语语言治疗服务的人数将非常庞大。随着公众对于言语治疗的认知逐步提高，“贵人语迟”、“不说话只是内向”等观念逐渐减少，越来越多言语障碍的孩子和成人及时接受言语治疗，追逐“与世界轻松对话”的梦想。同时，中国对特殊教育群体的重视程度提高，投入更多的资源到这个行业。大环境为中国言语治疗的发展提供了前所未有的良好机遇。

口吃者通过线下的口吃纠正班来纠正口吃，其费用少则三五千，多则一两万，

而对于口吃者本身大多数是弱势群体，其收入普遍较低，这大大阻碍了口吃者主动接收治疗。不仅如此，对于正规且大规模的口吃纠正班，大多数存在于大都市，不少的口吃者需要长途跋涉去千里外才能得到良好的治疗。这路程之远不仅加重了口吃者的经济压力，而且极大的加重了他们的心理与身体疲劳，从而治疗效果大打折扣。

基于以上口吃治疗遇到的种种困难，本系统充分发挥移动端的便利性，使得口吃者能够在 App 和相关治疗仪器上，随时随地进行训练，体验到良好的治疗效果。

1.4 平台的亮点和创新之处

1.4.1 概念创新

本平台搭载的 iSpeak 口吃辅助治疗系统是目前国内首个致力于口吃治疗的智能化移动端解决方案，具有一定的前瞻性与创新性。

口吃的治疗，原本是需要专业的医师与医疗机械的辅助，而高昂的价格、大量的时间精力投入需求以及良莠不齐的行业现状令很多口吃患者望而却步。在很多经济水平相对落后的地区，专业的口吃治疗机构更是不见踪影。我们通过把治疗方式移动端化，实体化，极大降低了使用和入门的门槛，使得生活中更多的口吃者在任何时刻、任何地点都可以得到及时有效的治疗，从而减轻他们因口吃带来的生活、心理等更方面的压力。

在任何学习过程中都存在一个共同的痛苦源，在心理学上叫做“心力萎顿”。屡屡受挫而导致的“习得性无助感”、逃避惩罚而形成的避免失败动机、过度强化使学习者失去学习兴趣等都是其形成的原因。这往往也是造成口吃习惯难以戒除、

导致口吃治疗失败的重要诱因。针对这种情况，我们在辅助治疗的移动端添加了社区圈子，进步历史记录等功能，有助于口吃者克服心理障碍，激发改造自己的潜力与动力。

1.4.2 市场创新

中国拥有 1300 万口吃患者，按照国际通行的需求标准，每 10 万人中需配备 20 名言语治疗师，按此标准，中国内地至少需要有 26 万名言语治疗师，而目前中国的言语治疗师队伍不足 1 万人。这一缺口无疑是口吃治疗庞大的市场的有力证明。

就目前的市场调研情况来看，国家目前没有对我国的口吃行业进行任何专门的监管，几乎任何人都可以从事这个领域。由于进入门槛太低，又缺乏相应督导，导致了目前从业人员的整体素质较低，水平较差，乱象较多。目前的口吃行业，负能量太多，问题不少，行业发展不太健康。

有些机构甚至还采取了“先矫正后收费”的模式，这种模式本身没有好坏，但这种无保障的霸王条款破坏了整个行业的收费模式与生态，从长远来讲，其实这是一种慢性自杀行为。

高昂的治疗价格、丛生的收费乱象，潜在地为本项目开发的移动端口吃辅助治疗平台提供了一定的市场机遇。相比之下，移动端产品价格较低，训练开展方便，性价极高，给他们带来的效用会远大于其付出的成本。

开发团队依托高校科研团队，有着较为雄厚的技术研发实力。机构产品打破口吃智能辅助行业的空白，技术含量较高，对于潜在竞争者具有一定的技术壁垒，在一定程度上抑制了同行竞争程度。

1.4.3 技术创新

平台建立了口吃词句的评价与分析数学模型、基于内容的语料库偏好推荐算法，使用基于定时采样和词级别分析的语音识别解析方案，开发了流畅发音训练与评分功能、阻塞音节分析功能、针对训练语料库推荐功能、社区经验分享与交友功能等特色功能。

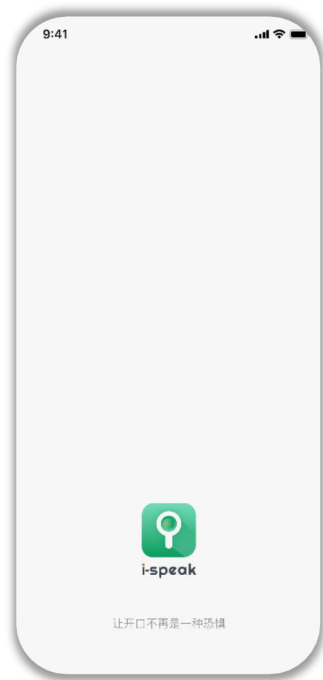
我们改进距离编辑算法，全盘考虑口吃发生的三种不同情境，使得评价具有足够的合理性与健壮性。为了给口吃者提供更加个性化的矫正方案，我们在项目中加入了基于内容推荐算法的语料库选择模型，建立了推荐语句库，通过统计方法按照训练模块进行推荐。我们还通过音节级别的词句分析，使用聚类算法将音节受阻的分析结果与推荐语句建立连接，进而推荐一个语句类别，将训练分析结果持久化。

二、特色功能概览及界面交互设计

2.1 基础功能简介

这里我们主要以 iOS 设备为例展示我们的基础交互面板，其他硬件载体效果如智能互动机器人的交互界面设计与移动端相似，效果详见演示视频。我们以用户友好、方便操作为设计理念，添加了一定的动态效果增强用户体验效果。

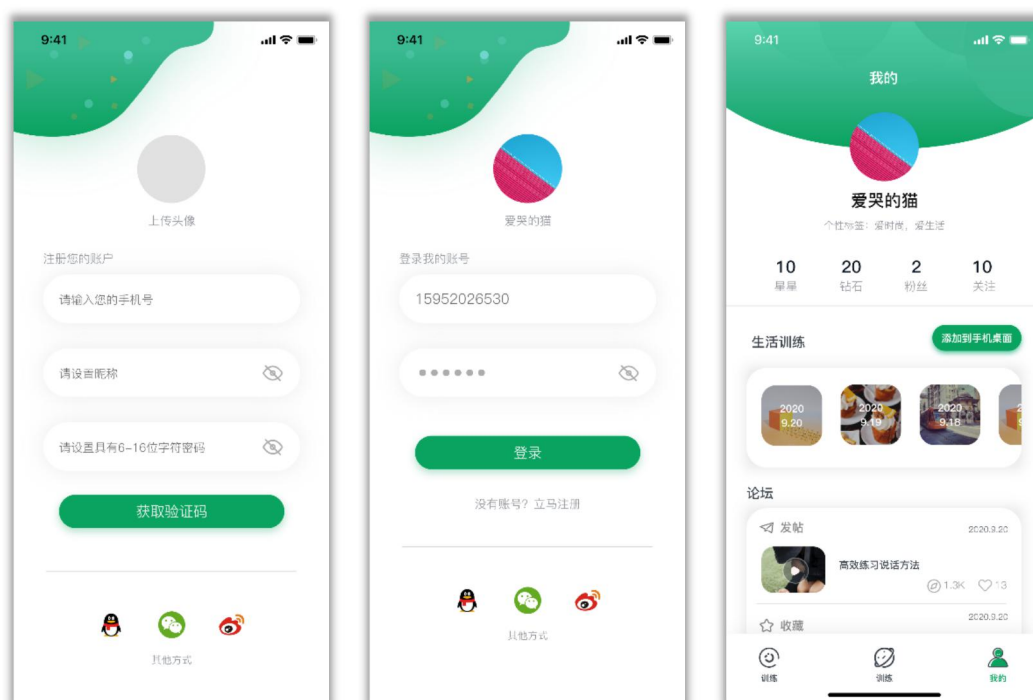
启动加载界面与 logo 设计：



基于 iOS13 新特性的 Widget 桌面小组件布局效果如下图所示,使用者无需
打开 App 即可开展生活场景练习:



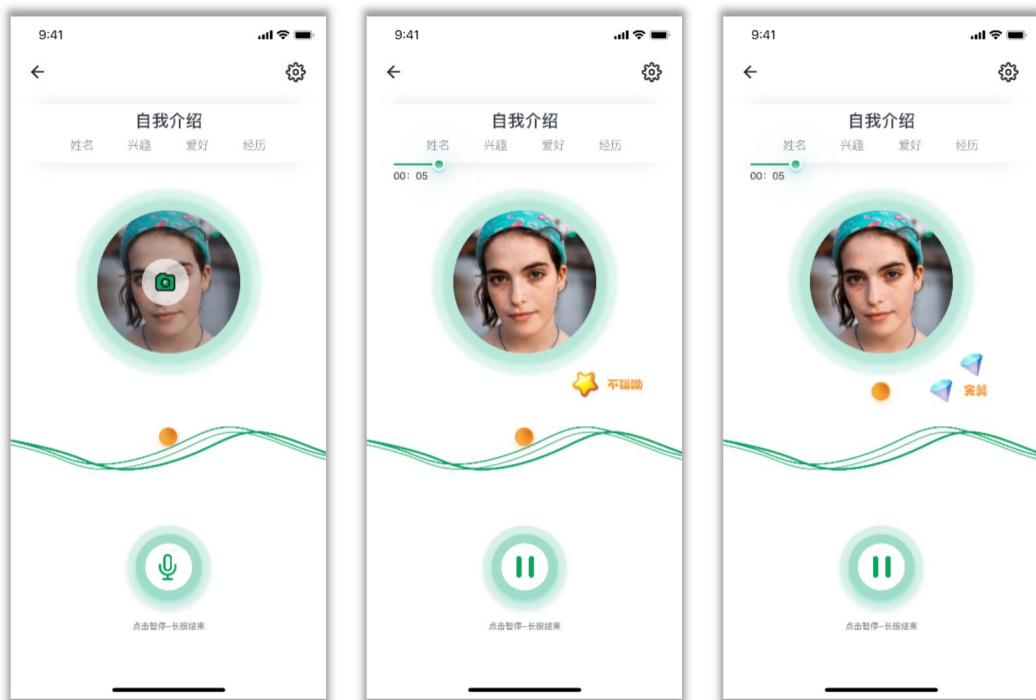
登录界面、注册界面、个人基本信息修改界面设计如下图所示，使用者通过手机号绑定注册账号，可设置昵称、性别、签名等个人信息：



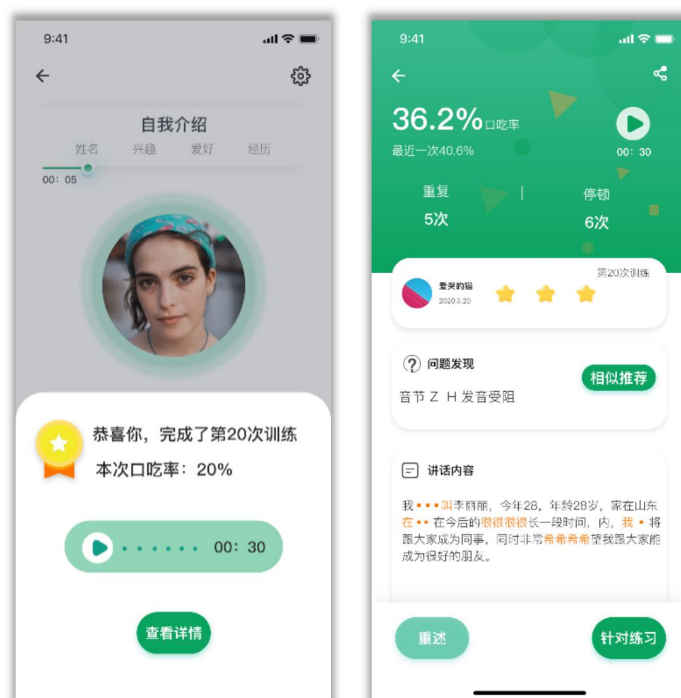
2.2 针对口吃语句的科学评分功能

发音检测，科学评分。流式语音听写使得平台可以对用户的发音作出迅速反馈，帮助口吃者通过有节奏的讲话训练，逐渐改变说话习惯。

录音检测界面具有丰富的动态用户交互功能，其 UI 界面设计如下。使用者点击录音按钮即可开始录音，再次点击即可暂停，长按录音按钮，即可结束录音，进入到评分界面；录音过程中，平台根据后台评分模型的定时反馈会给予用户一定的鼓励，通过小动画进行提示，并奖励应用内虚拟货币：



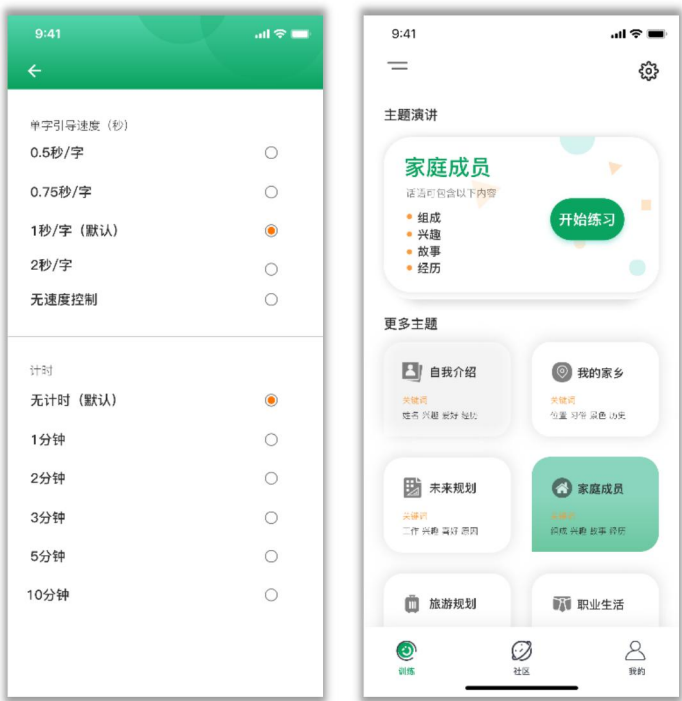
评分 UI 界面设计如下图所示，评分包括重复词句次数、较长卡顿次数以及总口吃率等基本分析数据，并会根据后台音节解析的结果，对用户重复与卡顿的词语和音节进行二次统计分析：



2.3 包含丰富语料模块的针对训练功能

针对训练，高效练习。对于口吃患者评分较低的词语或语句，进行针对性地推荐练习，更高效率地矫正口吃障碍。多种训练模式，可调节跟读语速，帮助患者快速找到适合自己的训练节奏。

应用内置调节引导参数功能，UI 界面设计如下图所示。用户根据自身口吃程度，调节跟读和引导速率，并设置计时选项，选择最适合自己的个性化训练模式：



根据用户发音发音的薄弱环节，我们使用基于内容的推荐算法来进行高效率的针对训练：



2.4 基于历史记录与音节解析的分析推荐功能

记录历史，进步点滴。APP 记录每个用户的评分历史。用户可以查看自己的进步过程，培养进步的成就感和继续训练的动力。

历史界面提供话题重训功能，方便用户查看进步点滴，有助于口吃者克服心理障碍，激发改造自己的潜力与动力，基于音节解析的分析模式，有助于用户发现自身发音的薄弱之处所在，进一步建立战胜口吃习惯的自信：



2.5 论坛交友与经验分享的社交功能

社区交友，走出阴影。建立口吃患者用户的专属社区，没有人有资格嘲笑其他人。社区内可以进行经验分享、提供随机话题模拟谈话情景等。平等相处，结下真心朋友，共同走出口吃障碍的阴影。

论坛版块 UI 界面设计如下，用户可以发布音频、图片、文字等动态，分享自己的经验与心情，并可以对其他用户的分享进行赠送应用内虚拟货币等互动：



社区也会定期推荐高质量和高浏览量的文章给用户，将有效的经验分享传播：



三、核心技术路线

2.1 基于定时采样和词级别分析的语音识别解析方案

对口吃者说话的结果进行语音识别是一个基础性问题。为此我们设计了基于定时采样和词级别分析的语音识别解析方案，对口吃者说话的结果进行识别和转换。

科大讯飞股份有限公司在语音技术领域是基础研究时间最长、资产规模最大、历届评测成绩最好、专业人才最多及市场占有率最高的公司，其智能语音核心技术代表了世界的最高水平。本平台嵌入了科大讯飞公司开发的流式语音识别 SDK，具有较好的识别率。

首先，设定识别间隙 ϵ ，为正常人的吐一个字需要的时间，通过相关资料的检索，我们在平台中使用时设定的识别间隙为 0.3 秒。识别过程中我们采用基于时间等距取样的方式对语音识别的结果进行采样，初始构造空字符串，如果相邻两次识别的结果并未发生变化则表示使用者并未吐字，此时在识别结果字符串中追加一个空格作为未识别出汉字的标识，如果想领两次识别的结果发生了变化则将新说出的字追加到识别结果字符串中。

其次，需要将识别结果字符串转化成正常语句字符串。平台设计的算法为：去除字符串中的空格，按照字长 1 从头至尾扫描字符串，忽略“我是”、“我叫”等指示词后面的 2 至 3 个字符的重复，即考虑到名字中重复的字，保证转化算法的健壮性。按照字长 2 以同样方式进行扫描，扫除口吃者吐词重复的现象，通过上述步骤即得到转换字符串。

2.2 基于改进后的距离编辑算法的评价与分析数学模型

对于口吃者说话口吃率的评价是一个综合性的问题，在 iSpesk 平台中基于口吃者所说的话对其进行口吃率评价是一个重点问题，在此我们通过改进的距离编辑算法对其进行计算。

首先考虑如下三种情况：

(1) 口吃者 1 说了一句包含 10 个字的话，到了第 8 个字开始出现口吃情况，口吃者 2 同样说了一句 10 个字的话，到了第 3 个字就开始出现口吃情况，算法中应当给出口吃者 2 的口吃率更加高。

因为口吃者开始说话时容易产生心理不适应，口吃者在随着训练次数增多的情况下能够改善口吃心理上的不适，此时开始说话时不容易产生口吃，在随后如果产生口吃则与最初练习时刻的口吃率相比应该有所提升。

(2) 口吃者只能吐一个字，如第 88 界奥斯卡最佳真人短片《口吃》中的主角，存在严重的元音发音受阻，吐字经常是“I…I…”，甚至是“I……”，所以算法中应当考虑完全无法吐字的情形。

(3) 口吃者说话中存在许多重合，比如“我是…是…”，相对于正常的话“我是”多了时间间隙以及重复的吐字。算法中应当考虑重复或者拖长的音节。

我们得到识别出的原字符串和转化成的正常字符串后，假定口吃者所说字符串为 S ，转换字符串为 T ，依据上述情况三，基于编辑距离算法对口吃程度进行初步衡量。编辑距离是指由原字符串 S 变化到目标字符串 T 所需最少的编辑操作次数。这里编辑操作的含义是：对字符串指定位置的单个字符进行插入、删除、替换的操作。这种对于单个字符的操作，称之为元操作。求解使用经典的 *Levenshtein* 算法。

假定 $S = s_1 s_2 s_3 \dots s_m$, $T = t_1 t_2 t_3 \dots t_n$, 建立 $S \rightarrow T$ 的匹配矩阵 $\Sigma_{(m+1) \cdot (n+1)} = \{d_{ij}\}_{(i,j)}$,

其中 d_{ij} 为字符串 $S' = s_1 s_2 s_3 \dots s_i$ 与字符串 $T' = t_1 t_2 t_3 \dots t_j$ 的编辑距离, 该编辑距离的计算本质上是一个最优化问题, 通过动态规划算法可以高效求解, 转移方程如下所示:

$$d_{ij} = \begin{cases} i & j = 0 \\ j & i = 0 \\ \min(d_{i-1,j-1}, d_{i,j-1}, d_{i-1,j}) + a_{ij} & i, j > 0 \end{cases}$$

其中 a_{ij} 的计算方式如下:

$$a_{ij} = \begin{cases} 0 & s_i = t_j \\ 1 & s_i \neq t_j \end{cases} \quad (i = 1 \dots m \ \& \ j = 1 \dots n)$$

令 ld 为 S 和 T 的编辑距离, 即 $ld = d_{mn}$ 在情况三中, 可用如下两种方式计算字符串的相似性:

$$Sim^* = 1 - \frac{ld}{\max(m, n)}$$

$$Sim^{**} = 1 - \frac{ld}{m + n}$$

对应的口吃率分别是 $1 - Sim^*$ 和 $1 - Sim^{**}$ 。

引入两个字符串 $S = s_1 s_2 s_3 \dots s_m$ 和 $T = t_1 t_2 t_3 \dots t_n$ 的最长公共字符串概念:

取 S 的子串 $S' = s_{i_1} s_{i_2} s_{i_3} \dots s_{i_k}$ 与 T 的子串 $T' = t_{j_1} t_{j_2} t_{j_3} \dots t_{j_k}$ 使得 $s_{i_r} = t_{j_r} (\forall r \ 1 \leq r \leq k)$,

则 $lcs = \max(k)$, 该最优化问题存在严格的求解算法, 使用动态规划算法求解方式如下所示:

$$lcs_{ij} = \max \begin{cases} lcs_{i-1,j} \\ lcs_{i,j-1} \\ lcs_{i-1,j-1} + 1 \text{ if } s_i = t_j \end{cases}$$

下面对编辑距离算法进行优化, 结合两个字符串的最长公共字符串长度 lcs ,

考虑情况 1，加入两个字符串第一次失配的位置参数 $\delta(1 \leq \delta \leq |S|)$ ，字符串相似度计算方式为：

$$Sim = \frac{lcs}{ld + lcs} \cdot \frac{\delta}{|S|}$$

遂得到口吃率 θ 为：

$$\theta = 1 - \frac{lcs}{ld + lcs} \cdot \frac{\delta}{|S|}$$

我们引入 $(|S| - \delta)/|S|$ 的原因是前驱字符对相似度的影响始终大于后继字符对相似度的影响。观察上述函数，容易发现在情景 1 的情况下说话的人先发生口吃比晚发生口吃的口吃率高，并且如果完全没有口吃现象，则口吃率为 0。

考虑到正常人的说话停顿以及一些特殊情况，需要对口吃率设定阈值进行修正，假定设定的阈值为 Δ ，可认为设定在 10%-15%之间，则口吃率的修正公式为：

$$\theta^* = \theta \cdot \mathbb{I}(\theta > \Delta)$$

其中 $\mathbb{I}(\cdot)$ 为指示函数，当条件为真时为 1，否则为 0。

考虑情况 2，对口吃者只能吐一个元音的情况，检查初始字符串中连续检测到相同字符串的次数，记为 L ，设定阈值 Δ_1 ，超过该阈值则直接标定口吃率为重度 $\alpha = 100\%$ ，修正公式：

$$\theta^* = (\theta \cdot \mathbb{I}(L < \Delta_1) + \mathbb{I}(L \geq \Delta_1) \cdot 100\%) \cdot \mathbb{I}(\theta > \Delta)$$

通过算法分析容易得到上述算法的时间复杂度为 $\Theta(m \cdot n)$ ，其中 m 和 n 分别是原字符串和转化字符串的长度。一般在三十秒内正常人的吐字不超过 90 字，可知算法可在较短时间内得到口吃率。

2.3 基于内容推荐算法的语料库选择模型

本平台为了给口吃者提供更加个性化的矫正方案，在项目中加入了基于内容推荐算法的语料库选择模型。

2.3.1 基于模块主题的推荐算法

推荐系统是通过历史记录进行针对性推荐的系统，包括信息获取、数据挖掘与针对性推荐三个方面。iSpeak 平台中保存了用户的历史数据，故对于单用户的信息获取并不困难。通过历史的评分记录以及发音受阻记录，可以得到训练的六个模块的熟练程度。熟练程度与模块训练的平均口吃率呈负相关，故六个模块（自我介绍、我的家乡、未来规划、家庭成员、旅游规划、职业生涯）的熟练度计算方式如下：

$$\phi_i = \frac{1}{\theta + \varepsilon} (i = 1, 2, 3, 4, 5, 6)$$
$$\varepsilon = 10^{-7}$$

将上述熟练度经过标准化处理映射到 $[0, 1]$ 区间，变换方法为：

$$\phi'_i = \frac{\phi_i - \min\{\phi_i\}}{\max\{\phi_i\} - \min\{\phi_i\}}$$

其次，由于熟练度为 $[0, 1]$ 之间的数，将上述熟练度经过如下转换得到单个模块的不熟练程度：

$$\phi''_i = 1 - \phi'_i$$

然后，按照统计方法，可以按照概率确定推荐的模块，每个模块被推荐的概率为：

$$p_i = \frac{\phi''_i}{\sum_i \phi''_i}$$

最终，被推荐模块编号为：

$$index = \underset{i}{argmax} p_i (i = 1, 2, 3, 4, 5, 6)$$

2.3.2 基于受阻音节分析的语料推荐算法

根据研究资料和现有的记录，按照发音部位可以将声母分为双唇音 (b、p、m)、唇齿音 (f)、舌尖前音 (z、c、s)、舌尖中音 (d、t、n、l)、舌尖后音 (zh、ch、sh、r)、舌面音 (j、q、x)、舌根音 (g、k、h) 这七个类别。按照发音方式可以分为塞音 (b、p、d、t、g、k)、塞擦音 (z、c、zh、ch、j、q)、擦音 (f、h、s、sh、r、x)、鼻音 (m、n)、边音 (l) 这五个方面。

分析口吃用户的重复与拖长声母或元音，即可得出用户发音受阻的音节，通过该音节进行上述归类即可得到用户发音的主要薄弱部位以及主要发音受阻方式。

对于口吃用户，基于统计方法对其受阻音节进行分析并推荐个性化的训练方案对纠正其口吃具有十分重要的意义。

定义基于发音声母的特征向量 $\gamma = \{y_1, y_2, \dots, y_{26}\}$ ， $y_i (i = 1, 2, \dots, 26)$ 按照如下方式进行计算：

$$y_i = \begin{cases} 1 & y_i \text{ 为发音受阻音节} \\ 0 & y_i \text{ 不是发音受阻音节} \end{cases}$$

下面使用 $K - means^+$ 对该分析结果样本所属的类别进行鉴别。

$K - means^+$ 聚类算法属于一种无监督式学习算法，通过统计方法将各个样本点归到所属的类别中。假设数据集 D 被该算法分成 K 个簇 c_1, c_2, \dots, c_K ，簇的质心点分别是 $\mu_1, \mu_2, \dots, \mu_K$ ，数据集中的样本记为 x 。

$K - means^+$ 的目标是使得一下簇内的误差平方和 (SSE) 最小。簇内的误差平方和最小意味着簇内样本集的相似度较高。

$$E = \sum_{k=1}^K \sum_{x \in c_k} \|x - \mu_k\|^2$$

$K - means^+$ 算法的步骤如下:

1. 产生初始簇质心点

(1) 创建空集合 S ，用于存储簇的 K 个质心点。

(2) 从数据集 D 中随机抽取一个实例，作为第一个簇的质心点 μ_1 ，添加到 S 中。

(3) 对于数据集 D 中的每个实例 x_i ，计算到 S 内个簇质心点的距离的平方，将其中的最小值记为 x_i 到集合 S 的距离的平方。计算方式如下所示：

$$d(x_i, S)^2 = \min_{\mu_j \in S} \|x_i - \mu_j\|^2$$

(4) 从 $D - S$ 中以如下加权概率分布随机产生下一个簇的质心点 μ_i ，并添加到集合 S 中。这样做使得初始的 K 个质心点以比较高的概率保持较远的距离，使得聚类效果比朴素随机产生初始簇中心的算法更好。

$$P(x_i) = \frac{d(x_i, S)^2}{\sum_j d(x_j, S)^2}$$

2. 对于数据集 D 中的每个实例 x_i ，分别计算到各簇质心点的距离，将 x_i 划分到与其距离最近的质心点所代表的簇。

3. 所有实例划分到各簇之后，各簇使用簇内实例重新计算质心点。设定 $center(c_k, j)$ 为簇 c_k 质心点的第 j 个维度的值，则计算方式如下所示：

$$center(c_k, j) = \frac{\sum_{i \in c_k} x_{ij}}{|c_k|}$$

4. 重复步骤 2 和步骤 3，直到各簇质心点不再变化（或变化很小）或迭代到指定次数。

对于语料库中的任意句子，抽离出其中所有的声母和元音，构造如上所示的特征向量 γ 。对特征向量进行聚类，指定类别为 7 种，通过语料库中的句子对 7 个类别的簇中心进行拟合，最终得到七个簇中心的中心点，记为 $\gamma_i^* (i=1, 2, 3, 4, 5, 6, 7)$ ，将语料库中的句子按照离中心点的距离分别归属到 7 个类别中的一个中，并添加标签。

一旦对口吃者的受阻音节进行评价完毕之后，构造相应的特征向量，按照离簇中心的距离进行升序排序，将其归入距离其最近的类别中即可得到口吃者本次口吃结果分析所推荐的语句类别，从该类别中进行随机抽样即可得到推荐语句。该算法支持一次训练多次推荐，具有一定的健壮性。

2.4 技术路线可行性分析

此前，新加坡综合医院及口吃诊所的语言治疗师使用双语评估英语和普通话双语患者的口吃情况，包括口吃病史的采集、双语语言样本的收集和评估，并且用两种语言分析言语口吃行为、次要肢体行为和对交流的负面反应等。治疗师使用 Lidcombe 行为数据语言分析每种语言的口吃比例，口吃严重程度及口吃行为类型，分析后确定相应的治疗干预。

首都医科大学附属北京儿童医院儿童保健中心主任医师刘春阳老师对口吃因素的研究中，被试是因口吃来北京儿童医院就诊，每说 100 个字至少有 5 次说话不流利，持续时间超过 6 个月，并且没有其他神经问题的患儿。但研究者没有具体描述如何测量每 100 个字有 5 次说话中断。

也有一些研究试图使评估方法更为客观。方芳提出了一种基于语音识别技术的口吃评估方法，以避免治疗师在评估过程中的主观性。她建立了人工神经网络

和马尔可夫模型，实现了自动口吃识别。使用马尔可夫模型将口吃样本分为 4 类，阻塞语音、重复语音、拉长语音和流利语音，准确率达到 75%。台湾的一项研究中研究者也有类似想法，他们尝试使用端点检测和动态时间扭曲技术自动识别口吃语音。

考虑到中国言语治疗师数量有限，基于计算机评估可能是评估口吃语音的好方法。这些想法被我们融入了技术路线的设计方案中，这保证了技术路线实现效果的可行性。

四、商业推广计划

4.1 市场调研及结果分析

◇目前我国口吃群体基数大，并保持一定的增长

口吃者即语言残疾患者，口吃是一种言语障碍，表现为言语频繁地与正常流利的人在频率和强度上不同、且非自愿的重复（语音，音节，单词或短语）、停顿、拖长打断。根据 2019 年全国口吃者群体抽样调查，截止 2019 年底，我国有口吃语言障碍的人数达到了 1300 万，其中男女比例是六比一。而中、小学生中的口吃人数更为惊人，大约占学生总人数的百分之二左右。并且总体而言，男性受到的影响比女性多两倍至五倍。但是，关于这种性别比例偏高的根本原因，目前知之甚少。大多数口吃始于儿童早期，研究表明 2.5% 的 5 岁以下儿童患有口吃。从刚开始口吃的儿童中可以看出，口吃的男孩和女孩数量相当。不过，性别比例似乎随着孩子的成长而扩大：在学龄前儿童中，口吃的男孩人数比口吃的女孩比例大约为二比一或更少的男孩少。这一比例在一年级时扩大到三比一，在五年级时扩大到五比一，因为女孩的康复率更高。由于早期恢复率高（约

65-75%)，口吃的总体患病率通常被认为约为 1%。

◇对口吃的观念

全世界有超过 7000 万人受到发展性口吃的影响。重要的是要与公众，科学和医学界以及那些口吃的人建立联系，以提高对口吃的认识。从这个简短的角度来看，我们认为要教育，倡导和传播有关口吃的意识，我们需要榜样，支持和机会。

但口吃者在社会中往往容易受到歧视。比如遭受嘲笑，学业成绩不佳，对将来的工作、婚恋，以及未来走向社会各种各样的影响。在国内，其现象多表现为多数口吃者的自闭、抑郁等自我情感障碍现象。尤其是在童年时期，口吃幼儿容易被同龄人排挤，歧视甚至霸凌。口吃者总是要接受别人抛来异样的目光，而这无疑严重的影响到了他们的正常生活。在现代，口吃的污名化已经十分严重，口吃即结巴，是一种言语障碍的病症。但却成为了日常中的贬义词，加之社会对于这一部分群体的关注力度较小。在美国法律中，明确将口吃者归类到残疾人一类。而在我国，口吃者并不属于残疾人。这就导致了并无一个强有力的官方机构合理有序，有针对性地管制这一领域，这导致对于口吃者的社会关注度较低。

4.2 目标服务群体分析

本团队走访南京特殊教育师范学校，口吃者贴吧以及口吃者聚集区展开相关调研。在对调研的 236 名对象中，94.2%的人希望能更好地融入到社会中，渴望与健康人进行无障碍交流。84.4%的人对本项目的口吃治疗及测评系统表现出兴趣。48.3%的人有直接的购买欲望。口吃治疗平台对于口吃者来说虽然不属于必需品，但随着自身对社会融入需求的增大，机构产品有着很高的欢迎度。在产

品价格方面，多数人希望产品定价能在每月 20 元左右，其中选择定价为 15 元的人数较多。



据全国残疾人抽样调查和有关机构统计，截止到 2019 年年底，我国有口吃语言障碍人数已达 1300 万人，同时我国每年口吃者的发病率达到了 1%。同时我国口吃行业发展缓慢，除了少数公益性组织外，社会上口吃治疗机构力量极其薄弱。同时培养一个专业的口吃治疗人才成本也十分昂贵，想要培养出一个合格的口吃治疗人才至少也需要两到三年的时间。并且由于我国目前的口吃治疗者的工作待遇并不是特别好，这也间接导致了优秀口吃人才的流失。由此可见，在我国如此庞大的口吃者群体下，我国口吃治疗行业缺口巨大。

在口吃者治疗市场存在巨大空缺的现状下，针对口吃者渴望正常交流的需求，一方面，本产品旨在帮助口吃者群体进行日常的交流，避免他们在公共场所面临空有需求却无对策的窘境。另一方面，本产品的目标市场主要集中在口吃者群体口吃者研究领域的参考辅助用品，可以帮助口吃治疗培训机构培养口吃治疗人才，大大降低培养口吃治疗人才的成本。对于口吃人群，口吃治疗社区系统便应运而生。

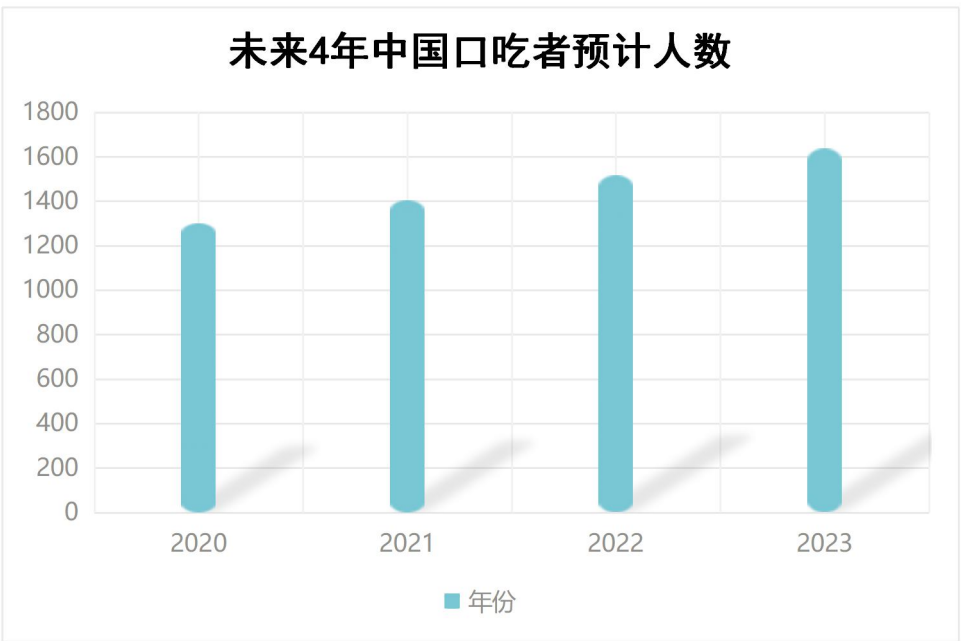
我们将市场细分为三部分，针对口吃者群体，我们推出口吃治疗训练实现口吃者的发音纠正和信心练习，找到自己对于某些发音点的缺点与克服技巧；针对

因口吃影响到社交的群体，我们建议将推出口吃者交流社区。在社区中，口吃者群体可以分享自己的口吃经历，以及自己的克服经验。与同是口吃者的伙伴们相互交流与倾心探讨。最后一部分是利用我们的核心口吃率测评算法计算口吃率。这并不仅仅只是为口吃者推出的功能，对于一些要去演讲或者做答辩的人群。口吃率可以很好的反映出他们实时的语言表达情况并且给出量化的反馈。

4.3 市场推广预测

(1) 市场容量预测

我国有口吃语言障碍人数已达 1300 万人，大约占总人数 1%，同时我国每年口吃者的的增长率率达到了 0.08%。



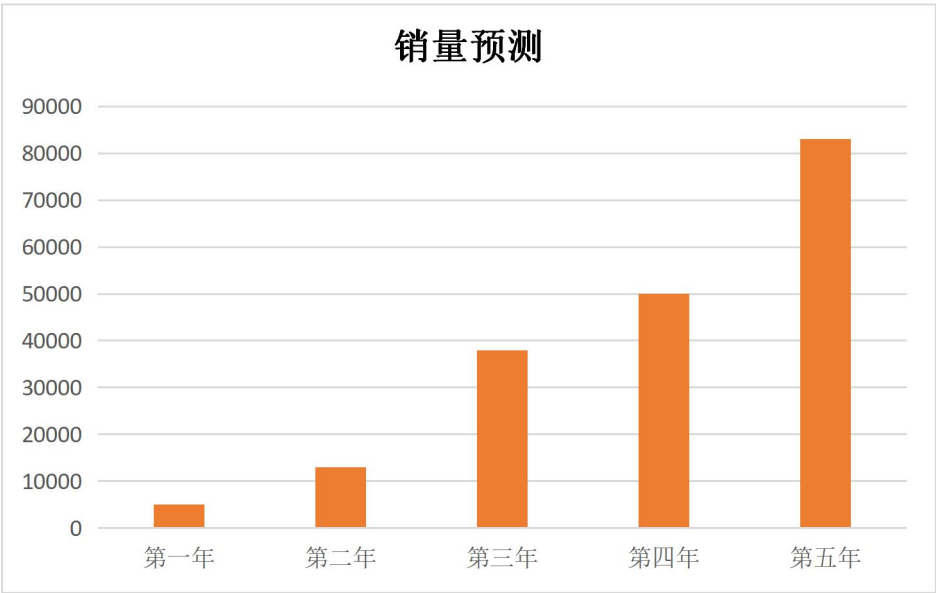
由上图可知，口吃者数在总人口数中有着相当可观的占比。基于对机构产品市场容量的预测，可以得出无论是口吃治疗翻译市场还是需要进行语言流畅度测评的市场均有着良好的发展前景。巨大的市场容量背后是对口吃治疗行业的巨大

需求。本团队将抓住机遇，致力于将 ispeak 口吃治疗平台推向社会，在此基础上研究开发，努力开拓市场。

(2) 产品销量预测

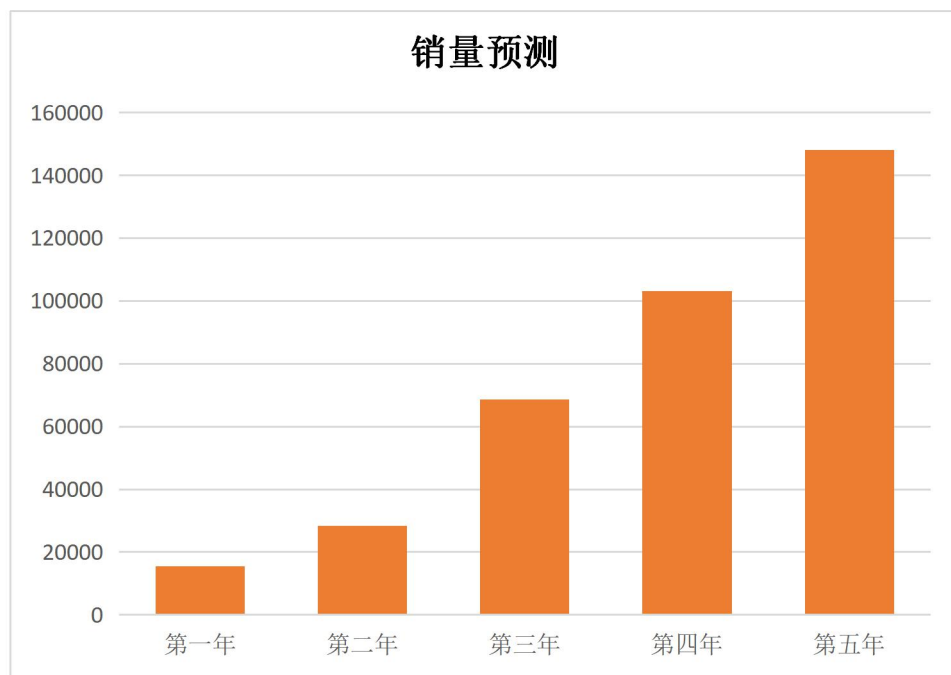
由于口吃治疗软件市场目前没有成品落地，因此进入壁垒不高，本团队要抓住机遇，以高性价比的产品迅速占领市场，提高产品的市场占有率。考虑到目前在口吃者口吃治疗工具的市场处于空白，基于前期的市场调研和相关估算，机构得出第一年在江苏地区销售 5000（人/月）的预测销量。

具体销量预测如下表：



第五年以后，本机构将目标客户重点定位于私人市场，并将线上线下有机结合，开拓全国市场。同时将扩大团队人数，开展全国范围内不同地区、不同场所的服务。

以下是全国销量预测：



4.4 SWOT 分析

经过充分的市场调研和比较，本团队的口吃治疗系统在国内市场上未遇到同类产品，对抢占市场先机，扩大市场规模具有先天优势。从长期来看，随着我国口吃者人群数量的不断扩大，该产品有良好的市场前景。但从短期来看，如何快速地进入市场，将机构产品推广出去，需要我机构制定完善而有效的战略计划，因此，我们通过以下的 SWOT 模型分析我机构的优势、劣势，面临的机会与威胁，同时制定机构的战略。

4.4.1 Strength：优势

(1) 人力资源优势：

机构不仅有专业的技术团队，更有着一批既具有洞察力，又有创造力的团队，团队成员为大学生，不仅有着扎实的知识基础和专业能力，同时创业、服务热情高涨。

(2) 技术优势：

机构依托高校科研团队，有着较为雄厚的技术研发实力。机构产品核心技术有大学实验室同学自主研发，打破口吃智能辅助行业的空白，技术含量较高，对于潜在竞争者具有一定的技术壁垒，具有广阔的市场前景。

(3) 产品优势：

产品针对口吃者人士一类特殊群体，成本较低，且可以较好地符合不同类型口吃者人士的需要、具有很好的价格优势，性价比较高，给他们带来的效用会远大于其付出的成本。

4.4.2 Weakness：劣势

(1) 资金相对缺乏，市场推广难度大

团队在起步时期由于没有形成产品的品牌优势，并且风险较大，融资相对困难，初始资金相对缺乏。同时由于在创立前期还没建立必要的关系网络，市场开发这一环节较为薄弱，这就对我们的营销环节提出了较高的要求。

(2) 经验不足，知名度较低

在具有社会成功创业者的经验传授与专业老师的创业指导的支持下，本团队主要由学生组成。作为学生初创团队，虽然有着跨学科的专业知识，但相较社会人士而言缺乏实战经验下有效运营和决策能力。由于前期产品市场、服务市场尚未打开，因此这需要我们团队更加注重实践经历，在摸索中前进，提升自己的能力共同致力于机构未来的发展。

4.4.3 Opportunity：机遇

(1) 市场需求增长强劲

通过前期的市场调研，可以看出很多口吃者士，特别是青少年群体对本机构的无障碍沟通系统很感兴趣，一旦产品投入市场，会有很大的市场需求。

(2) 关爱口吃者士意识普及

随着社会不断进步及社会人群素质的不断提高，人们关爱口吃者士的意识不断增强，这有利于他们社会融入感的提升和社会地位的提高。本机构产品在促进口吃者社会融合方面起到很好的辅助作用。

(3) 国家及地方政策支持

国家从“十二五”规划以来就一直重视残疾人等语言障碍事业的发展，并把残疾人早日实现小康生活纳入到工作日程中来，同时地方政府也在积极打造残疾人便利生活新模式。

(4) 创业孵化园政策支持

南京理工大学创业孵化园对创业大学生提供了优质，高效，便捷的创业服务，客观上为大学生创业提供了坚实的后盾。

(5) 国内处于智能设备发展高峰期

当前的时代是人工智能的时代，国家大力推进人工智能产业的发展，智能产品在近几年呈井喷式增长，这推动了智能技术的革新，同时也反作用与智能设备的更新换代。

4.4.4 Threat: 威胁

消费者认知不足

许多口吃者容易生活在极度自卑的生活状态下，导致对于许多新鲜事物出现抵触和难以接受被突然关心的感情。因此，对于让口吃者体验本产品是一个难点，同时也决定着本产品的市场评价。

五、结语：未来遐思

在未来一段时间，完全治愈重度口吃患者的任务亟难实现，但在对社会弱势群体关注度日渐上行的今天，这一梦想会越来越快地变为现实。

对于轻度或中度口吃患者，这款口吃辅助治疗系统提供的平台将是一个很好的训练工具，患者可以通过自我努力摆脱口吃阴影；对于重度口吃患者，这款辅助治疗平台或许也可以成为医生、患者的好帮手，医生通过平台数据跟踪患者康复进度，避免症状复发。

采用语音技术矫正口吃，在国内外都处于前期研究阶段，所能利用的研究成果较少。如何合成符合实际语音的自然度和流利度都较好的口吃矫正语音是人工智能在医疗上非常有价值的应用，同时也是语音合成技术进一步成熟完善的表现，并为其他疾病的计算机智能治疗开辟了道路。

项目组全体同学坚信，总有一天，所有人都能自信流利地表达自己。我们热切地期盼那一天的到来。