# 12345市民服务热线智能工单系统

### 1、现状及目标

#### 1.1 背景

上海12345市民服务热线是直接面向市民的综合政务服务一线窗口，代表了市热线办、全市140多家市政管理单位及公共事业机构。市民热线话务员良好的服务质量，直接向市民展现了政府的服务形象，体现了上海政务服务高品质。

12345市民热线自2012年开通运行以来，主体业务是受理市民来电，完成市民诉求的受理、转派、回访，与市民进行沟通互动的主要渠道是传统语音，在服务渠道方面，提供了电话语音、传真方式，受理市民诉求。围绕着为市民诉求解答、受理、反馈的闭环业务，热线中心形成了一套成熟的业务支撑平台。

目前，12345热线业务系统已建成话务系统、IVR系统、坐席系统、大屏幕系统、诉求单系统、知识库系统、考试培训系统、质检系统、统计报表系统、热线服务运营管理支撑系统等一系列的系统工具，为业务运营提供了一套功能全面、性能稳定的业务支撑系统平台。

为进一步提升服务渠道“运营智慧化”的服务能力，希望能够通过智能化手段，实现市民来访需求的快速记录，工单内容的智能分类及派发，关键热点问题的统计分析，市民情绪识别，以及工单处理的质量检测。达到加快现场话务员服务响应能效，提升整体服务质量与市民感知，提高热线整体运营数据分析的智能化程度，优化管理的效果。

热线主要智能化应用规划包括基于语音识别、机器学习、深度学习、知识图谱、情绪识别以及智能预质检方法实现对市民通话表达诉求的快速识别、诉求工单的智能化生成及处理、实时服务质量监督、智能化运营管理分析。同时后续考虑开展基于智能问答的自助人机交互功能，将人工语音咨询逐步向自助咨询引流，提升服务效率和服务感知。

12345市民服务热线目前日均呼入量为25000通，而接通率为92%。电话诉求日均受理量已达到20000+件，话务员每日人均受理量达到60件。随着12345对市民服务定位的不断提升；对诉求受理范围与处理效率的不断提高要求，话务员的服务受理压力越来越大。

目前话务员在每次受理市民诉求时，需要花费时间去理解和分析整理市民电话口述过程中所提出的诉求内容，并需要大量时间去进行工单填写、分类和承办单位的转派协调。这部分工作非常耗费时间人力，导致服务效率不高。

#### 1.2 数据

**一、工单数据**

工单数据为客服接通电话后，人工总结市民的整体诉求，并记录的数据。

*例如：市民来电咨询：已办理过居住证的外省市户籍人员是否可以在本市办理港澳通行证、所需材料、办理地址、办结时限以及工作时间。（市民要求发送短信）*

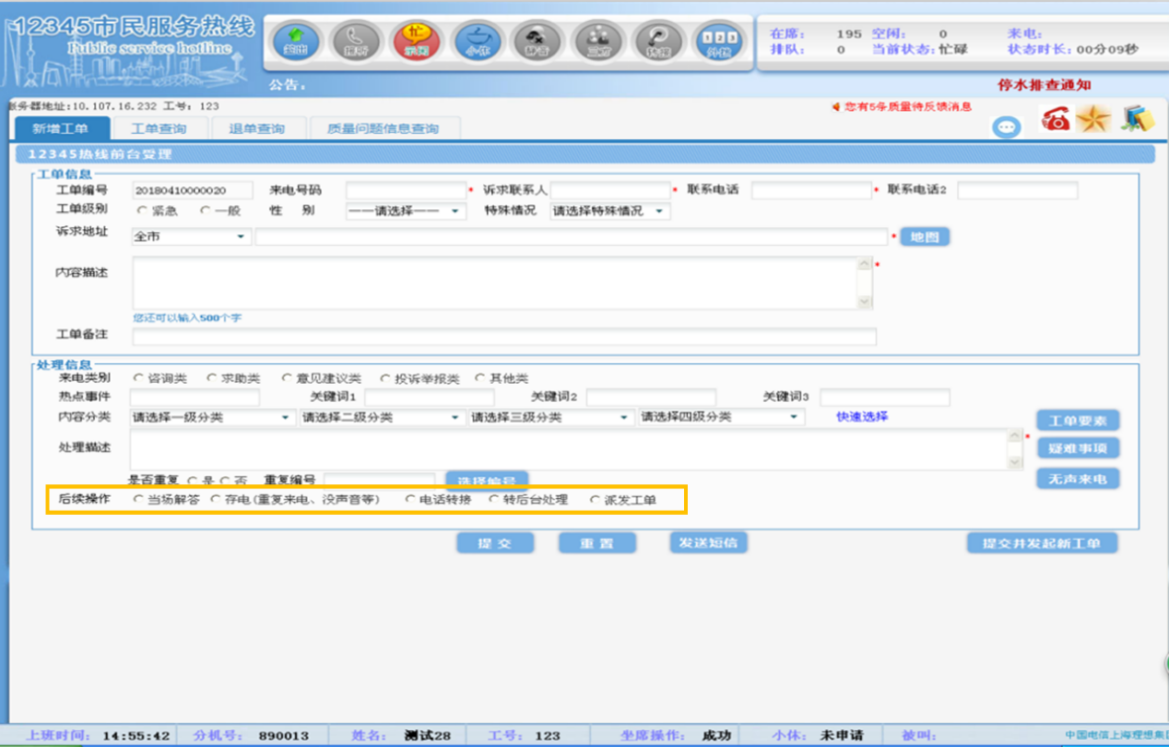
**二、语音转文本数据**

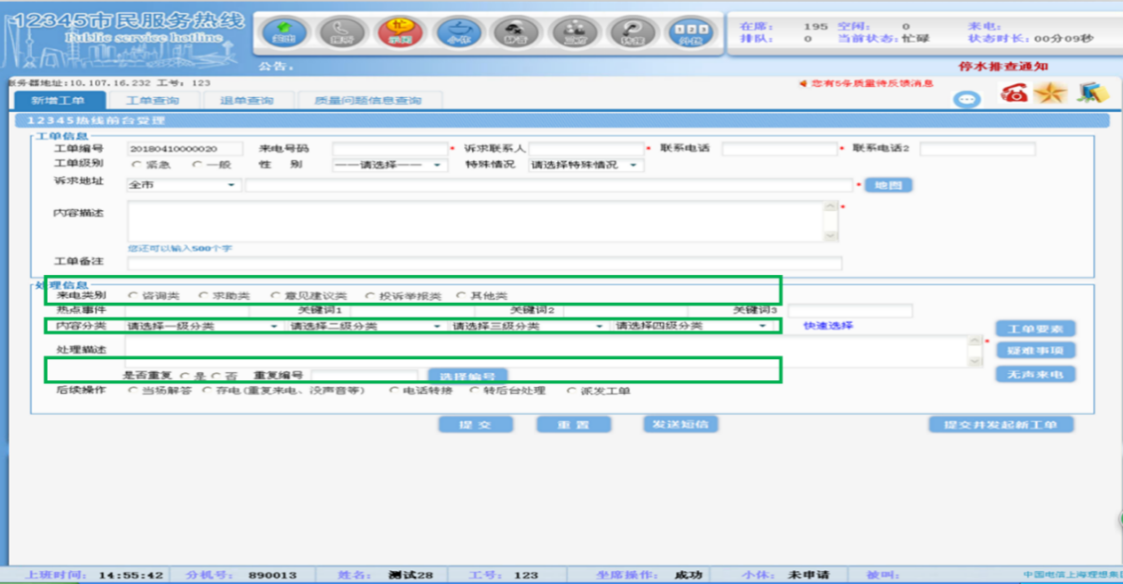
语音转文本数据为客服接通电话后，通过语音转文本引擎，将语音数据直接转译为文本数据，并保存。其中红色为客服语音文本数据，蓝色为市民语音文本数据。

*例如：,工号2566，为您服务,，您好，请问有什么可以帮您,喂，你好，我想了解一下，就是上海市现在垃圾分类的举报方式是什么,，呃我们热线帮您反映的话，也可以帮您记录反映，噢，是这样，我想先了解一下，因为的话我发现我们小区的话最近还是有很大问题,，你把哪些情况,由城镇垃圾分类，有问题的，我们热线就可以帮您记录，反映，需要的话您可以直接我们热线，我帮您记录反映,，好的，然后的话我想知道就是行政流程是什么，就是我举报之后，然后的话你这边的话就是噢，安排人过来就是检查，对吧,？按照热线的办理要求和流程，我们是帮您记录反映到管理部门，要管理部门进行处理，他们是不是会再派人到现场，我们电话当中无法为您直接判断,，嗯,，好的，我知道了，然后的话处罚的话是按照标准处罚对吧,？具体的处理情况的话，您可以,在此之前我们热线有结果的话，我们可以将结果告知到您,，好的谢谢，我知道了，我知道了，谢谢,，不用谢，有需要的话您再致电我们热线先生,，好的，再见,，如果您没有其他夫妻交，稍后请对我的服务做出评价，感谢您的来电再见,。好的,。*

其中语音转文本数据，客服语音转文本的准确率为90%，而市民语音转文本的准确率仅为70%，文本质量较差，在建立工单分类模型时，准确率较低，因此后续工作都围绕工单数据展开。

#### 1.3 目标

为提升服务能效，希望能够通过人工智能相关技术为12345市民服务热线工单系统提供智能化填写功能。实现对工单数据、工单分类、派发流程的自动填写。

工单智能分类的准确率最终能达到80%以上。工单智能派发要求较高，工单派发单位要求准确率达到90%以上。

### 2、存在问题

目前工单数据在进行工单分类时，主要存在下面几个主要问题：

#### 2.1 分类体系大

工单类目体系共包括4层，第一层的分类数量是9，第二层的分类数量是63，第三层的分类数量是256，第四层的分类数量是700+。例，安全监管类的类目体系如图所示。



#### 2.2 数据样本不均衡

工单数量严重不均衡。工单大类和小类的数据比值最大达到428。

|  |  |  |
| --- | --- | --- |
| **第一级分类名称** | **数量** | **比例** |
| 安全监管类 | 1105 | 14.73333 |
| 公安政法类 | 16465 | 219.5333 |
| 公用事业类 | 3801 | 50.68 |
| 建设交通类 | 32103 | 428.04 |
| 经济综合类 | 2054 | 27.38667 |
| 科教文卫类 | 4195 | 55.93333 |
| 其他类 | 6353 | 84.70667 |
| 社会管理类 | 14367 | 191.56 |
| 社会团体类 | 75 | 1 |

#### 2.3 分类类目之间易混淆

工单本身的类目体系之间有较多的类区别不明显，客服打标容易出现错误。不同的客服的理解标准不同，容易出现打标混淆，例如：

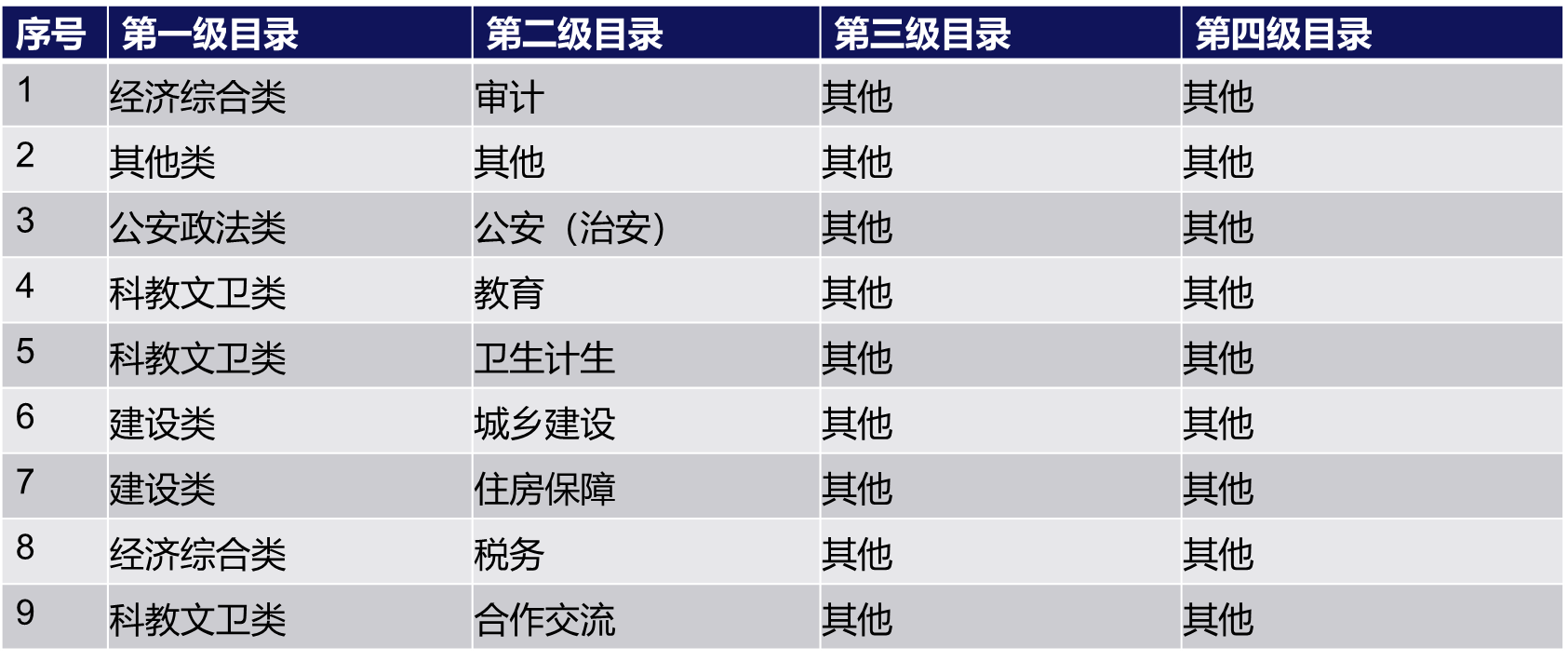
#### 2.4 单条工单多个诉求

市民来电反映时，如果涉及到多个事件或诉求，需要建立多个工单，而这些工单之间的重复性较高，难以分类。例如：



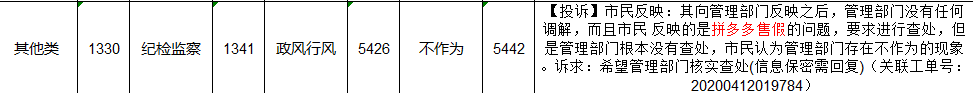
#### 2.5 四级分类中的（\*\*-其他-其他）数据

\*\*-其他-其他类数据比较分散，而且与具体的分类有数据重叠。在进行分类时，很难被分到具体的其他-其他类，但是该类在工单类目中又是不可或缺的。而\*\*-其他-其他类共有60个，其工单数据量占总工单数据量的9%。



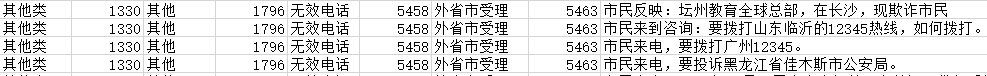
#### 2.6 涉及态度相关类难以分类

市民来电反映时，如果涉及到事件投诉或工作人员的投诉，工单需要对具体事件进行描述，会误导工单分类。例如：



#### 2.7 存在较难进行数据建模的类

工单分类类目中，存在一些咨询相关的类，而咨询类工单较短，关键词少且咨询的范围分散，较难建模。例如：





#### 2.8 工单派发准确率低

工单派发方向有140左右，并非所有的数据都涉及到工单派发，但派发单位对派发工单的准确率要求较高。

#### 2.9 其他

客服本身分类错误、工单描述不精确等干扰问题。