

亲历

记者与高校团队、民间人士为乡音存档

教AI学温州话

■ 本报记者 张源 项锐

“天不怕,地不怕,就怕温州人说温州话。”这句民间俗语,道尽了温州方言的独特与难懂。作为吴语区存古度高、语音复杂的代表方言,温州话素有“古汉语活化石”之称。如今随着普通话普及,温州不少年轻人都不太会说温州话,守护乡音迫在眉睫。

今年春天,“温州方言系统化采集工程”启动。该工程由浙大城市学院承办,浙江大学文学院学术指导,阿里巴巴AI数据团队、国家语言文字推广基地(温州大学)联合支持,他们集结了1.5万名温州市民参与录制,累计采集有效语料超6400小时。团队不只为方言建“声音博物馆”,更要让AI听懂、说准地道温州话。

近日,我们走进浙大城市学院和温州城乡巷弄,跟着高校技术团队与民间爱好者一起教AI学温州话。

“和人学外语一样,不能只背单词”

上午9时,浙大城市学院杭州语言服务协同创新中心里,温州话此起彼伏,这里是“温州方言系统化采集工程”的核心审核现场。推开门,项目首席专家、浙大城市学院外国语学院教授姜淑珍紧盯电脑屏幕,点开一段段音频,耳朵敏锐地捕捉着每一个音节。

我们刚要打招呼,她的手机先响了。电话那头是一位温州阿伯:“我讲了一辈子正宗温州话,能不能也给AI留一份?”

姜淑珍用温州话感谢他,并耐心解释,目前采集入口已关闭,后续若有需求会主动联系他。挂断电话,她转头笑着对我们说:“这几天每天都有十通这样的电话,七八十岁的老人、在外工作的温州人甚至新温州人都想加入。”

她调出系统后台的问卷给我们看:温州家庭里,晚辈与长辈方言交流比例超六成,长辈对晚辈全程讲温州话的仅剩13%。她坦言,00后大多能听不能讲,地道发音率偏低,不少俚语、古音连中年人都未必全掌握,“再不做抢救性采集,以后想听正宗温州话,只能找档案了”,这正是启动大规模采集的深层原因。

过往方言保护,多由少数资深发言人在录音室录制固定词表与语句,成品

封存档案馆,普通人接触不到,也无法活态传承。因语料形态单一,脱离真实场景,无法适配AI训练对鲜活语料的核心需求。本次采集彻底打破这一模式,核心收录带有人情味的“活语言”。题库中不只有固定文本,也有自由话题,其中不乏日常交流、就医购物等真实生活场景。既有专业人士参加,也向温州民众全面开放,并兼容周边区县。

“AI学方言,和人学外语一样,不能只背单词。”姜淑珍打了个比方,“你得让它‘浸’在真实语境里,听不同年龄的人在不同的场景下怎么说。只有放进生活,AI才能学到最鲜活的表达。”

我们当场掏出手机试了试采集测试端:填写基本信息后,进入5句试音。屏幕上,每句都配齐方言汉字、普通话翻译,生僻字还有读音提示,瞥一眼音标就能想起。只要逐字朗读,口音大体地道即可通过,随后进入正式录制。全程手机操作,几分钟到几小时皆可。

“原定采集2000小时语料,短短十几天我们就收得6400小时优质语料,大大超出预期。”姜淑珍介绍,为兼顾语料标准与方言多样性,采集内容覆盖鹿城、瓯海、龙湾、永嘉、瑞安等核心片区口音。

音文对齐率不低于95%,否则AI就会学偏

采集工作圆满收官,真正的考验才刚开始。如何将上万人录制的口素质材,转化为AI可学的“标准教材”?带着这个疑问,午后,我们在姜淑珍的指导下,当起了临时“方言审核员”,亲历了三道严苛的审核关卡。

第一道关,是人工初审。作为语料入库的首道质控关口,审核员必须逐条细听、逐字核对。我们点开音频文件,仔细核对,耳边不时传来团队成员细致地判定:“这句说得挺地道,通过”“末尾掺杂了普通话,不通过”。

有两段听感相近的音频,判定结果却截然不同。为什么?“仔细听。”姜淑珍笑着回放音频并解释,“温州话里‘看’和‘吃’都读ci,但声调完全不同。”‘吃’为先降后升的曲折调,而‘看’为降调。”她反复播放三遍逐字对比,我们才终于听出了其中的关键差异。

温州话保留了中古汉语声调体系,有复杂的连读变调和轻重音规则,本地人凭语感能轻松分辨,但对AI而言却是难题。“同一个词语,只要轻重读法不同,



记者与新温州人一起学习温州方言。

本报记者 周旭辉 摄



姜淑珍教授教记者张源(左)审核采集到的音频。

本报记者 周旭辉 摄

意思就天差地别。”她说。

初审后,第二道关便是异常值算法智能校验。系统通过音长、波形、频率等特征类比,自动识别与基准样本偏差大的异常音频,规避人工主观偏差,提升入库精度。

“这段鹿城林阿姨录制的文本是‘我有听灵清’(我没听清楚),她顺口在最前面加了一个虚词‘句’(ha)字,所以算法自动把这句标为不通过。”姜淑珍指着屏

幕讲解。

“日常这么说没问题啊?”我们问,她点头道,平时这样表达确实自然又地道。但方言AI训练的核心指标是音文对齐率不低于95%。若有偏差,AI会学得错误对应关系,导致模型识别、合成出现系统性偏差。

我们审核了20多条音频,发现不合格的样本五花八门。有人漏读了虚词,有人掺杂了普通话,还有人脱离文本按

日常口语习惯自行发挥……但凡音文没有严格对齐的样本,一律不入库。

最后一道关,是双轨制复核,为前两道校准兜底。由技术合作方独立质检组抽检核验音义对齐精度,项目专家团队同步专业校准。截至目前,音频验收全部通过,字准确率约98%。经过三关严格的筛选,最终入库的语料既符合AI训练

的标准化要求,又保留年龄、性别、语速、地域等自然语音差异。这种“斤斤计较”,为的是让AI学到的温州话,能和街头巷尾的真实乡音对上号。

AI方言一旦落地,就能派上大用场

实验室里的6400小时数据,最终会应用到哪里?200多公里外的温州方言学习班,正好解答了这个问题。

晚上7时,温州瓯海区新桥街道党群服务中心的活动室里,十多位学员正齐声练基础发音。这里是温州话爱好者叶一戈的方言学习班,已经开办了2年,场场爆满。我们找到位置坐下,跟着大家一起学讲温州话。

课堂上,叶一戈带着大家反复练习最基础也最难发的音。温州话没有翘舌音,也不分前鼻音,完整保留着普通

早已消失的清浊音。我们跟着念“医”和“夜”,前者发音喉咽要收紧,后者声带有震动感。细微差别却对应完全不同的意思,跟读十来次我们才勉强摸到门道,周围学员也个个眉头紧锁。

后排一位穿白T恤的男学员听得格外专注,笔记本上记得十分详细。他是温医大附二院关节骨病科主治医师李振兴,河北人,虽然来温州七八年,但对温州话依然没有那么精通。“门诊七成以上老人只会说温州话。”他无奈地说,像膝盖、腰椎、关节等诊疗关键词,温州话说法和普通话差十万八千里,猜都没法猜。

“腿弯这里酸”“腰骨落不起”……老人用方言描述,他常常只能靠手势猜部位和病情。“疼痛位置、受伤机制、发病时间,哪一句听岔了,都可能影响诊断。”为此他下定决心来学习温州话,想把关键表述啃下来。

“如果AI真能学会温州话,诊室里老人说方言,医生实时看文字转写,普通话方言互译,能解决我现在遇到的大难题。”他感叹道。

教室里的其他人同样需求明确。温州大学东北籍周老师要和华侨深度交流,不懂方言常卡壳;基层干部希望靠方言拉近与群众的距离;温州女婿、媳妇想融入家庭;还有本地年轻人因家人无人教,特意来补课。这些真实场景,无一不在印证,AI方言一旦落地,就能派上大用场。

“AI要学的不只是发音,更是语言背后真实鲜活的表达与文化逻辑。”叶一戈上课时举过一个极具温州特色的细节:温州人不说“鸭舌”,而称“鸭近”。因为“舌”与表亏本的“折”同音,而“近”等同于“赚”,寓意财源广进。一句方言,就是一套独有的生活逻辑与地域文化密码。

方言数字化与AI应用探索还将持续进行。目前,杭州、温州试点已经完成采集,宁波同步启动,西安、长沙等全国各地的试点也在稳步推进。

AI学方言,终究在于“用”。要让老人能用方言叫外卖、挂号、打车,让新来的人不用再靠猜和比划。只有当技术落到这些实处,乡音才能真正“活”下去。



张源



项锐

普陀“共富方舟”驶入桃花岛,记者跟船体验——漂洋过海,送优质服务上岛

■ 本报记者 黄宁璐

在舟山普陀,有一艘特殊的船,不载游客,只送“健康”与“幸福”,名字叫“共富方舟”。

作为浙江省海岛数量最多的县区,普陀共有大小岛屿7435个,其中住人岛455个,长期以来面临公共服务覆盖不均、资源下沉困难的现实挑战。近年来,普陀区锚定基本公共服务一体化改革,创新打造“共富方舟·惠民行动”,将优质资源打包上船,送到海岛百姓身边。目前,该行动已累计开展67次,覆盖所有住人岛。

从倾听民声,到统筹资源,再到漂洋过海送服务……近日,我跟随工作人员,亲历了“共富方舟”驶入桃花岛茅山庙的全过程。

笔记本上记下岛上百姓细碎的期盼

距离“共富方舟”起航还有一周,我乘船登上桃花岛,跟随茅山村党总支书记虞哲将走村入户,摸排需求。

“我们村地处偏远,住的多是六七十岁老人。‘共富方舟’每次来送的都是实实在在的便利,大伙儿格外珍惜。所以得先上门问问,大家最需要啥。”虞哲将边走边解释,“当然,光靠起航前这一趟不够。平时,村干部、网格员串门唠嗑时就在排摸需求了。两者结合,把服务送到百姓心坎上。”

田间地头,几位村民正忙活着,虞哲将扯开嗓子喊:“下周,区里‘共富方舟’又要来送服务嘞,想要啥,快来讲!”话音刚落,田里的村民纷纷直起腰,院前摘菜、门口闲聊的老人也围拢过来,你一言我一语,争先恐后地念叨起自己的心愿。

74岁的史夫华说:“这段时间膝关节老痛,蹲也蹲不下。去大医院做坐船坐车,碰上起风起雾停航还走不了。医生要是能到村里来,那是再好不过了。”

我主动当起虞哲的助手,掏出笔记本,写下:史夫华,关节痛,盼专科医生下乡。

“睡眠不好,经常头晕,眼睛一吹风就流泪。”71岁的胡美红感慨,自己不识字,去趟县医院得子女陪,耽误孩子工作。“平时小病小痛,要去乡镇卫生院,要么自己硬扛。现在专家来,想把些毛病让他们瞧瞧。”

“天气暖和了,想把头发剪短些。”“这个季节嘛,最需要的就是当季能下播的种子。”“要是能看场越剧,就更美了。”……村民们七嘴八舌,我一项项记下,心里也跟着触动:这些细碎的期盼里,藏着岛上百姓的不便和无奈。

一上午,我们从村头走到村尾,见着村民就停下来聊几句,笔记本上记了20多个需求。虞哲将说:“现在老人们的孩子们基本都搬去岛外了,借着问需求,也能看看他们过得好不好。”

回来后,我们把需求信息整理好,发给了统筹“共富方舟·惠民行动”的普陀区委社会工作部。

跟着方舟行动,省力又高效

我马不停蹄地从桃花岛赶回普陀本

岛,跟着普陀区委社会工作部社会工作人员与志愿服务科科长刘佳侃一头扎进资源统筹工作。只见他将活动时间、地点和收集上来的百姓需求,发到“共富方舟·惠民行动联络群”里。

让我意外的是,信息刚发出,普陀区卫健局、区文化馆、区民政局等单位纷纷回复“参加”。一时间,“叮叮”声此起彼伏。

“最开始可不是这样。”刘佳侃回忆,2022年7月“共富方舟”启航时还只是卫健部门的医疗服务船,2024年升级后交由新成立的社工部统筹,“起初协调服务资源时,有的单位觉得这是‘多出来的活’。可后来他们发现,跟着方舟行动,既省力又高效——大件物资统一运,人员统一走;服务聚到一起,来的百姓也多了,大家的成就感更强了。现在,越来越多的单位主动要求加入。”

接下来两天,我协助他逐一联系报名单位,反复确认人员数量、标识牌和设备清单,与医院核对参与科室与医生名单,确保主要需求得到落实。30多个通话记录、5页A4纸草稿——我体验

后才知,看似简单的工作背后有多繁琐。

准备工作还没有结束。行动前一天,我跟着刘佳侃和几名工作人员搭乘“共富方舟”前往桃花岛,一同上船的还有DR车、冲击波理疗仪、B超机等医疗设备以及“菜篮子”流动惠民车等。“老人们习惯早起,我们得提前把设备运上岛、接好电源、搭好帐篷,第二天工作人员一到,就能直接开诊。”刘佳侃说。

我们赶到茅山村党群服务中心时,镇里和村里的人已把帐篷运到现场。“人员畅通是首位。”经验丰富的刘佳侃用脚步丈量场地空间后说道,“医疗是大头,帐篷分医疗区和其他便民服务区两块。配药窗口人流最大,得放在宽敞处;理发需求也大,且会有碎头发,找个不碍事的角落……”方案迅速定下,大家分头行动。

搭帐篷远比我想象中不易。篷布又沉又韧,撑杆卡扣总对不准,风一吹篷布就乱飘,刚支起一个角,另一个角又塌了……我手忙脚乱地拽布、插杆,

不一会儿就满头大汗了。“这已经是最舒服的时节了!”刘佳侃边说边过来帮助我,“每月至少两次,我们雷打不动送服务。最难熬的是夏天,在太阳下搭帐篷,干完活汗水都结成白花花的盐粒了。”

不知不觉夜幕降临,U字型蓝色帐篷区终于搭好,镇上电工确认所有设施设备都能正常通电,大家才安心回去休息。

忙得腰酸背痛,心里却暖洋洋的

次日清晨7时30分,茅山村党群服务中心就迎来不少老人。不一会儿,两辆大巴车缓缓停靠。一大早从普陀本岛赶来的近80名医护人员、志愿者快速下车,各就各位:医疗队接运医保设备、摆好诊疗器械;志愿者支起服务台、整理物资;越剧演员化妆备场……很快,海岛老人盼望已久的“共富方舟·惠民行动”就热气腾腾地开场了。

“阿婆,您是后发性白内障,得去医院做激光治疗。”眼科摊位前,医生仔



活动现场,医疗服务区围满了前来就诊的老百姓。

舟山市普陀区委社会工作部供图



记者(右)为老人查看药品说明。

通讯员 方碧蓉 摄

细检查后给出建议。“我之前做过白内障手术,最近视力又开始模糊。今天给医生看过,心里就有数了,这就去医院治。”74岁的杨友娣感激地说。医生递给杨阿婆一张医院联系卡——去年起普陀医院就为海岛老人开通了绿色通道,可电话预约专家号,到院后还有工作人员全程陪诊。

“5G+AI智能甲状腺检查”摊位前排起长队。医生陈厥一边做B超,一边登记信息。我主动上前帮他记录。“这套设备是医院去年引进的,增加了AI智能筛查功能,筛查更准更省时。”陈厥说,尽管拆装麻烦,他们仍坚持带来,让海岛老人用上先进的医疗设备。

普陀区“菜篮子”流动惠民车前同样热闹。村民李阿婆往篮子里装了好几样:“平时买菜得坐车去镇上,这里方便又实惠。”短短两小时,“菜篮子”车就卖出200多单。见工作人员忙不过来,我赶紧挽起袖子加入。俯身装菜、抬手递送,来回十几趟,腰背渐渐发酸。可看到老人们高兴的神情,顿时那点酸胀也就不觉得了。

一旁的茅山村文化礼堂里,越剧演员已穿戴整齐。锣鼓一响,水袖轻扬,《梁祝·十八相送》的唱腔婉转而起。台下老人们有的跟着节拍轻轻点头,有的眯着眼睛听得入神,嘴角不自觉上扬。农资发放摊位前,300份当季种子不到20分钟便一抢而空。现场还有理发磨刀、医保咨询、法律服务、药事服务等便民项目。

一上午,我哪里需要就往哪里补,手脚没停过,一番忙碌下来腰酸背痛,心里却暖洋洋的。看着那些老人脸上绽开的笑容,我忽然明白了“共富方舟”风雨无阻、漂洋过海的意义。



黄宁璐