

# 三十年枫红，华彩竞放

## ——访信研院语音和语言技术中心主任郑方

○周景萍 吴群慧



### 郑方

清华大学研究员，博士生导师，清华大学信息技术研究院副院长，语音和语言技术研究中心主任。北京得意音通技术有限责任公司创办人。曾获北京市高等教育教学成果一等奖、北京市科技进步二等奖和公安部科学技术三等奖。

在我们的生活中会有这样的情况：不该被访问的信息被访问了，不该被进入的系统被进入了，等等，这些都表明不安全的隐患随时存在。怎样解决这些问题呢？学界普遍认为由生物的肌体特征来判断肌体是最好的方式，这其中包括指纹、人脸、虹膜、声音。通过生物特征来判断，肌体有信息不易被窃取，不易被遗忘，使用成本低等优势。多选一的说话人辨认、说话人身份识别等即采用了这种判断机制。

郑方所在研究中心的三大研究领域即为语音识别、说话人识别和自然语言理解。语音和语言技术中心是一个跨计算机系和电子工程系的学科交叉平台，是由系的相关实验室等多个机构及人员组合而成。郑方和我们谈起，1986年他担任电子工程系实验室主任之初，曾提出启动相关研究的建议，即从基于规则和基于统计的两种语言理解技术出发，逐步学习并掌握口语对话系统和语言理解技术。实验室的方棣棠、吴文虎、李树青、蔡莲红、胡起秀等老师，认为应该集中力量进行基于我们所擅长的信号处理技术的语音处理技术，而不能分散力量启动基于他们所不熟悉的语言处理技术的口语对话系统研究。“他们并没有因为我是最年轻的老师就否定我的计划。经过几次讨论他们最终同意我的方案并给予大力支持。”

1998年，作为学校985项目的子项，在学校

和智能技术与系统国家重点实验室的支持下，语音实验室正式开展了口语对话系统的研究。经过9年时间，到2007年已经有了非常成熟的开发工具SDS、SDK的雏形。它脱离了常规的解决方法和解决思路，在“自然语言理解”、“对话管理”两方面能做很多应用。智能的搜索查询，例如与中国移动合作进行的垂直搜索研究就是中心的成果之一。

郑方及其研究团队在“语言理解”（人机对话）研究过程中发现，特定领域的语言理解比较容易实现，例如日常生活、餐饮、购物等。通过设计文法、语言文法描述、刻录，实现机器对人语言的理解。在他们与中国移动的合作项目中，只要拨特服号，衣食住行都可直接通过口语询问，语言被理解后，系统会给你发短信回答你的问题。著名的Google公司做的是水平搜索，而这种叫做垂直搜索，人不会被海量信息所淹没，能够更简单、更直接、更全面地搜索结果。

上世纪90年代末郑方做的是“语音识别”，2000年初启动“说话人识别”研究，并与公司合作付诸应用。这两项研究有相似也有不同，都是声音却个性不同。10年过去了，现在他们的研究已经处于国内领先地位，国家行业标准也是他们来牵头制定的。郑方及团队正在做着几个大项目，一个是和中国建设银行合作，通过声音认证，进行网上支付和电话支付。由于他们研制出的系统性能好，最终从多家竞争单位中胜出。他预计清华百年校庆之前能完成这个项目，这种识别方式将是金融第一例，也是一种全新的模式。另一个就是与中国移动合作的语音垂直搜索，预计广州亚运会时会使用。在“说话人识别”的研究中，他们还在防控反恐方面与公安部合作开发类似产品的应用。

凡是有人地方就会涉及语音与语言技术的应用。郑方设想以后不需要打字，只要说话就可以完成一篇文章。开车时只要通过叫出名字就可以拨通电话。他们很早就做了声控号码簿，里面

存有200个电话，只要说出名字它就能自动为用户拨号。快速准确地获取需要的信息将成为信息时代的基本特征。

当谈到信研院的产学研特色时，郑方道出他自己的理解：信研院的设立是要把信息学院好的成果通过学科交叉及大项目组织平台转化为成果。在郑方看来产学研绝不仅仅是个口号，一定要做到有机结合，有机互动。

信研院里的研究成果既要能用到实际生产上，还应当能把企业在应用中遇到的问题再返回来，建立回馈机制。信研院再针对实际问题进行研究、建模，在此过程中理论也会有所提高，研究才能真正得以深入。最后形成产学研有机的循环互动。理论与应用应该是一个交替上升的过程。

郑方提到，信研院每年与企业签约成立的联合机构有十几家，就是在做这样的尝试。校企双方共同解决问题，效果极好。有的企业已经连续签了三期，每期三年。这样做的好处，一是企业受益，二是信研院不断有新课题，三是提升教学质量，信研院的老师在参与教学的过程中，把实践中的问题与方案以及成果也都传达给了学生。郑方高兴地说，这样才越做越有味道。

他还谈到，近年来我国在语音识别、语言理解等领域都有较大发展，在理论和技术上不一定落后于发达国家。虽然国内和国外技术相比，国外提出理论较早，但我们学到了这些理论，从理论层面说，是站在他们肩膀上。但不是有了理论就能做出东西。周围环境、滋生平台等成熟的重要资源、环境保障是取得技术上突破进展不可或缺的支持。在国外，无论是国家政府还是民间都有较大支持，比如对课题组进行资金支持，组织



亚运城市名片启动仪式

课题组之间的评测活动等。

目前郑方等联系国际、国内八家单位形成中国语言资源联盟（CCC），建立了IDC语言数据资源，即语音识别、说话人识别数据库。他们正为丰富资源、建立平台做努力，也取得了一定进展。郑方相信如果能有更好的平台支持、国家导向，他们会做得更好。此外，他们还在努力克服的一个瓶颈是中文中有众多复杂的方言给语音识别等带来的难度。郑方说他们也有比国外超前的地方，像垂直语言搜索，国外就极少，只有一个美国的大学在做，还在研究之中。

语音和语言技术中心逐渐发展壮大，现在已经发展得很好。郑方说，我们一定会继续努力，争取在语音、语言领域取得更大的发展，在学术领域获得长足的进步，为清华百年献礼！

## 科研成果

### ■ 声纹识别技术 ■

声纹是生物特征之一。与其他生物特征相比，声纹更容易获取，采集设备成本低廉，而且蕴涵说话人的真实意图。声纹识别技术可广泛用于公安国安、金融交易、个性化服务等领域的身份识别。

2004年5月，信研院语音语言中心（CSLT）参与承担的“司法语音自动分析和鉴别系统的研制”课题，通过了公安部鉴定。2008年2月，CSLT参与承担的“通用声纹识别身份证

系统引擎的研制”项目通过了北京市科委的验收，专家组一致认为项目成果达到了国际先进水平。2008年3月，CSLT参与起草的《自动声纹识别（说话人识别）技术规范》标准由国家信息产业部正式颁布实施。2010年10月，CSLT参与研发的声纹识别系统经过国内某银行全面的功能及性能测试，已集成到电话银行系统中用于用户身份确认，并在深圳和武汉试点运行。