



一名小朋友在2019智博会主会场重庆国际博览中心参观

“干啥子哟？”“你瞅啥？瞅你咋地？”

在2019中国国际智能产业博览会(简称“智博会”)上,一款内嵌了最新人工智能方言识别技术的智能翻译机逗乐了来自天南海北的观众。

无论是四川话、东北方言,还是吴侬软语,都被科大讯飞翻译机逐字逐句转换成汉语文本,再实时翻译成英文。

此次参展的阿里巴巴人工智能实验室也开发了可识别四川方言的智能音箱“天猫精灵”,预计今年9月后正式发布,未来还会以智能音箱终端为基础,通过语音交互采样建立覆盖全国的动态方言数据库。

人工智能,正在帮助我们留下“多彩乡音”。

方言是中国语言文化的瑰宝,传承丰富的历史文化内涵;方言也是人们乡情的根系,具有丰富的人文价值。在我国,部分“小方言”在经济社会飞速发展的背景下有消失风险,方言的保护和留存极具紧迫性。

长期在中国进行汉语方言调研的韩国首尔大学中文系主任朴正九教授接受新华社记者采访时说,汉语方言使用人口多、地域广,在全球语言学研究占有非常重要的地位,汉语方言有助于研究人类语言的共性和个性。

早在2012年,科大讯飞就以粤语为突破口,开展方言识别技术攻关。2017年,科大讯飞输入法联合中国声谷发起“方言保护计划”,近3年来已收集超过86万条方言语音,支持23种方言语音输入,实现包括粤语、四川话、湖南话等11种方言的语音合成;微信小程序“我AI说方言”还方便公众随时随地上传方言,建设中国方言库。

阿里巴巴人工智能实验室今年也宣布,正式成立方言保护专项小组,投入1亿元人民币对汉语方言进行保护和开发。

科大讯飞输入法业务部副总经理、方言保护计划发起人李强军介绍,方言语料的采集、记录和归纳是方言识别的基础。以前,这项工作很大程

度依赖于调查者主观感知,而归纳完整的语音变化、进行句法和语义分析等工作难以单靠人力完成,所以有必要建立分属不同方言的数据库,利用人工智能系统地对方言文字、发音等进行整理。

“各种汉语方言在语法语音上的特征都不相同,很难被系统整理。用上人工智能技术,对方言保护和研究大有帮助。”朴正九说。

其实,外语语音识别与方言识别的人工智能训练方法是相通的。随着技术进步,可用统一的语音识别模型,导入各地方言等不同语料进行无监督训练。我国语言学界将现代汉语方言分为十大方言区,方言识别难度有所不同,落实到输入法产品中,识别准确率也略有差异。

“适配不同的方言识别,对整体语音识别技术也会是一种促进和优化。”阿里巴巴人工智能实验室高级产品专家张平介绍,做方言是为了让更多地域人群能更好地交流,同时也更多保护地域文化。“我们第一款先做四川方言,就是因为四川方言在中国方言中用户占比大”。

开发“方言版”语音识别,还有助于设计出更贴合消费者需求的智能语音交互产品。美国“风险投资节奏(Venturebeat.com)”网站数据显示,中国已超过美国成为全球最大的智能语音市场,2019年第一季度智能语音出货量占全球总出货量一半以上。

国外智能语音产品厂商也在努力开发能识别各种英语口音的版本。有趣的是,曾经有一位美国消费者投诉亚马逊公司的智能语音助手“亚历克萨(Alexa)”,因为它不能识别自己母亲的口音。

目前,亚马逊的“亚历克萨”、苹果的Siri和谷歌助手均可以识别标准美音、“印度味”英语和“中国味”英语。不过,据Vocalize.ai实验室测评,这三种产品识别中国口音英语的准确率最低。另外,它们都没有识别汉语方言的功能。

由此可见,独具中国特色的“方言版”语音识别产品还有助于中国企业开展差异化竞争。

据新华社

趣知道

世界各地最早的古人类都是谁

说起人类祖先,许多人印象中“辈分”最高的可能是著名的“人类老祖母”露西,但实际上在非洲发现的另一个古人类图迈的“年纪”约是露西的两倍。在中国,北京猿人闻名遐迩,但要论辈分也坐不上头把交椅。那么,世界各地最早的古人类都是谁呢?

新华社记者就此采访了中国科学院古脊椎动物与古人类研究所研究员吴秀杰,她介绍了迄今在世界各地发现的最早古人类。

非洲是学术界公认的“人类摇篮”。吴秀杰说:“迄今已知的200万年前的古人类化石,都是在非洲发现的。”其中,1974年在埃塞俄比亚发现的古人类化石露西,距今约350万年,曾长期被认为是最古老的人类,享有“人类老祖母”的美誉。

不过,后来在非洲又陆续发现了年代更久的古人类化石。如2000年在肯尼亚发现的干禧人,因发现时间而得名,年代距今约600万年。2001年在非洲中部乍得发现的图迈,“年纪”更是露西的约两倍,达到约700万年。图迈因此不仅成为“非洲冠军”,也夺得了迄今已知最早古人类的“全球总冠军”。

据介绍,在欧洲,约175万年前,德马尼西人生活在今天的格鲁吉亚一带。至于比较出名的尼安德特人,则并非以年代见长,而是因为他们可能与现代人的直接祖先有诸多互动而在当今引起广泛兴趣,他们生活的时间很靠近现代人,在约10万年至几万年前。

在亚洲,印度尼西亚爪哇岛上的爪哇人可能追溯到约180万年前,在中国发现的元谋人可追溯到约170万年前,近来的新研究将蓝田人的时间也推至约160万年前。由于现有技术测量误差等原因,上述年代在一百七八十年左右的古人类,可能还难以确定真实的先后。

据介绍,许多人从小就听说的北京猿人或北京人,近来研究显示年代在约77万年前,也就是说按“辈分”无法与元谋人和蓝田人相比。北京人的意义,既来自它在中国古人类体系中承上启下的地位,也来自它的独特经历——在1929年那个中国积贫积弱的年代,裴文中等中国考古人员发现了中国大地上的古人类化石,相关化石随后还经历了抗战期间失踪等备受关注的事件。

“世界上年代久远的古人类化石主要在非洲、欧洲和亚洲,在美洲和大洋洲迄今只发现了一两万年前的古人类化石。”吴秀杰说,这可能与古人类的生活范围有关,也可能与环境因素导致一些地方的古人类遗存难以形成化石并保存下来有关。

据新华社

科普

我国科学家发现水稻早熟高产新机制

水稻如何能既早熟又高产?我国多个科研团队合作发现,一个名为Ef-cd的基因可将水稻成熟期提早7天至20天,不造成产量损失甚至具有不同程度的增产效果。挖掘和利用该基因,将有力促进绿色超级稻品种培育的减肥增效需求。该研究成果日前在线发表于知名学术期刊《美国科学院院刊》。

在我国杂交水稻发展的早期阶段,“高产不早熟、早熟不高产”现象是一个重大难题。袁隆平院士团队和谢华安院士团队等经过研究攻关,培育出一批早熟高产的水稻新品种,但其分子生理机制仍未得到解析。

中国科学院遗传与发育生物学研究所储成才研究组和四川农业大学邓晓建研究组等单位合作,历经多年,图位克隆了Ef-cd基因。大规模组学分析表明,含Ef-cd基因的水稻,氮代谢、叶绿素代谢及光合作用相关基因表达显著增强。生理试验也证明,Ef-cd基因显著提高了水稻氮吸收能力及叶片光合作用效率。

上述结果表明,Ef-cd基因兼顾早熟和高产,增加水稻的氮肥利用效率和光合效率,具有资源高效利用的显著特征。研究团队表示,在当前水稻生产新形势下,该基因的挖掘和利用将有力促进绿色超级稻品种培育的减肥增效需求,对解决直播稻和粮经、粮菜、粮油连作稻、双季稻的早熟丰产以及亚种间杂交稻“超亲晚熟”等问题具有重要的应用潜力。

据新华社

浙大学者发明仿生修补液可令牙齿“长”出牙釉质

浙江大学化学系的唐睿康教授团队发明出一种仿生修补液,在牙釉质的缺损处滴上两滴,缺损表面在48小时内就能“长”出2.5微米晶体修复层,其成分、微观结构和力学性能与天然牙釉质几乎一致,并与原有组织无缝联结,浑然一体。相关论文于北京时间8月31日在线发表于《科学进展》杂志。

研究人员介绍说,牙釉质是人体中最硬的天然生物材料。这层包裹于牙齿表面的半透明的物质,厚度约为2毫米,其无机矿物含量高达96%。由于缺乏包括细胞在内的生物有机基质,因此无法再生。自恒牙长成的第一天起,牙釉质就在缓慢地消耗着,细菌酵解食物中的糖类物质释放出酸,以及酸性饮料都会加速它的消耗。一旦牙釉质的防线被突破,整颗牙就像失去了保护伞。

修复牙釉质,堪称是仿生领域一项最“硬”的挑战。常见的补牙材料,例如复合树脂、陶瓷和汞合金等,它们几乎发挥着“填料”的功能,适用于“大洞”修补,但对小裂却常常束手无策,与天然组织之间也不能完全结合。

“理想的修复方法,应该是材料、结构、力学性能三者的统一,而且能实现原位修复。”论文共同通讯作者、浙江大学化学系刘昭明博士说。

实验中,研究团队成员将富含磷酸钙团簇的溶液,用滴管滴在人工牙齿表面,随后将其放入到一个模拟口腔唾液环境的溶液中。“牙齿的表面首先形成了一个仿生矿化前沿。”唐睿康说,实验测量显示,48小时后,牙釉质“长”高了2到3微米,“也就是说,牙齿上长出了一种连续的材料,一个与原组织一模一样、完全结合的生物结构。”

研究还进一步测试了修复材料的力学性能,结果显示,长出来的人工牙釉质,其硬度和弹性模量与天然牙釉质的数值几乎相同。德国著名生物矿化学家、康斯坦兹大学赫尔穆特·科芬教授评价说,这项材料有望在临床上真正实现牙釉质的原位修复。

据新华社