

“皇冠上的明珠中最亮的那颗”

AI 何时才能与人类对答如流？

■本报记者 沈春蕾

为什么“小度”“小爱”等智能助理可以与人类进行对话和交流呢？为什么它们有时候也会“不知所云”“答非所问”呢？

清华大学计算机科学与技术系长聘副教授黄民烈告诉《中国科学报》：“如果说自然语言处理是人工智能(AI)‘皇冠上的明珠’，那么 AI 对话系统则是自然语言处理中最难、最核心的任务之一，是‘明珠中最亮的那颗’。”

日前，黄民烈联合 20 多位学者共同制定的全球首个《AI 对话系统分级定义》(以下简称《分级定义》)正式发布。

机器人也回答不了

图灵测试是常用的 AI 测试，测试者在不知情的情况下，向人类和 AI 系统随意提问。如果 AI 系统平均让每个测试者做出超过 30% 的错误，那么就认为 AI 系统通过了图灵测试，具有人类智能。

“起源于图灵测试的 AI 对话系统，是 AI 领域最重要的研究方向之一。”黄民烈说，当前，AI 对话系统被业内认为是衡量 AI 发展水平的重要因素，代表了 AI 的发展方向。

据了解，在工业应用领域，AI 对话系统已呈现出“爆炸式”增长的态势，其中，以“小度”“小爱”等为代表的智能助理，正广泛应用于个人助理、智能家居、智能汽车以及开放域闲聊产品中，与大众日常生活紧密相关。

然而，有业内人士指出，当前作为前沿技术的 AI 对话系统标准缺失，造成其在中出现水平参差不齐、评价体系不一的情况，导致业界因认知不一而误解 AI 交互水平，也引起了关于 AI 是否具有意识以及自然语言对话伦理和道德等方面的广泛讨论。

“《分级定义》的制定旨在衡量 AI 对话系统的能力水平，从而促进 AI 对话系统的进一步研究，同时为工业界应用提供参考。”黄民烈表示，考虑到 AI 对话系统任务繁多、评价维度多样、



黄民烈介绍《AI 对话系统分级定义》。受访者供图。

技术路线丰富，撰写小组在制定《分级定义》时仅关注完全由机器主导的对话系统，人机混合的对话系统不在考虑范围内。

人所掌握的知识是有限的，机器人也不例外。“一些简单的问题，对话系统可以回答，但一些开放边界的问题，对话系统受限于自身知识水平就很难回答。”华为诺亚方舟实验室语音语义首席科学家刘群向《中国科学报》举了一个例子，“有人在跟父母吵架后，问机器人该怎么办，机器人回答不了，这需要更多知识和场景的注入”。

黄民烈进一步解释道，为了在实际应用中发挥价值，《分级定义》的制定是从用户可感知以及可观察、可测量、可度量的角度出发，不考虑系统的具体技术实现方式，也不区分助理类任务、闲聊、知识对话等，均以“场景”进行表述。

衡量将有据可依

近年来，随着深度学习技术的不断发展，AI 对话系统已经从基于规则的第一代和以传统机

器学习为核心的第二代，发展到以大数据和大模型为显著特征的第三代，对话能力发生了革命性变化，在开放话题上展现了惊人的对话能力。不久前，谷歌研究员称 AI 已具备人格的讨论帖一度登上热搜。

黄民烈介绍，《分级定义》从自动对话能力、对话质量高低、单一 / 多个场景、跨场景的上下文依赖和自然切换能力、拟人化程度、主动和持续学习能力、多模态感知与表达能力等角度，将 AI 对话系统划分为 L0~L5 六个等级，等级越高，AI 对话系统水平越高。

小米集团技术委员会主席、AI 实验室主任王斌向《中国科学报》表示：“参照《分级定义》，当前工业界的 AI 对话系统水平最高已发展至 L2~L3 之间。”

但是，参加制定《分级定义》的专家学者也表示，AI 对话系统要达到更高级别还需要具备记忆、联想和推理等能力。

王斌在 AI 对话系统的研发过程中，时常感到难以评判所开发对话系统的水平。他期待《分级定义》发布后，AI 对话系统能力水平的衡量将有据可依——对于用户来说，能够

更多地关注、更清晰地理解 AI 对话系统；对行业来说，有助于企业明确研发方向，促进行业更规范发展。

AI 不会取代人

未来的 AI 对话系统会取代一些传统的职业工作吗？

“现在传媒界对 AI 比较恐慌，尤其是 AI 可以辅助写作。”北京师范大学新闻传播学院院长张洪忠告诉《中国科学报》，“AI 对传媒业而言是生产力的提升，不能将它妖魔化。以前是记者手写稿，后来是用电脑写稿，现在是 AI 可以辅助写作，后两者都是辅助工具，是生产力的提升，不会取代人。”

张洪忠一直关注社交机器人所带来的伦理挑战，他希望《分级定义》的发布为思考、讨论、制定法律规范提供有力参考，以规避 AI 对话可能产生的伦理问题。

刘群表示，即使是目前最好的 AI 对话系统，《分级定义》中 L4 等级提及的人设、人格、情感、观点也很难实现，“因为机器很难把握人类思想背后所隐含的背景知识、常识和逻辑等”。

刘群认为，《分级定义》能引发 AI 对话系统研究领域的思考和讨论，帮助研究人员更好地发现系统的问题并明确研究方向，其制定是一次大胆的尝试。

黄民烈表示，AI 对话系统最高级别的应用是复杂情感任务，《分级定义》的发布将有助于 AI 对话系统在情感任务中体现更高水平，从而促进 AI 在情感陪伴、心理健康、虚拟人、元宇宙等方面的应用，大大释放人力和物力成本，促进前沿科技走进大众日常生活。

此次《分级定义》的发布，是 AI 对话系统走向规范化、系统化发展的第一步。接下来，黄民烈将联合该领域相关研究机构及研究者开展白皮书的编纂，聚焦 AI 对话系统的发展历程，详细阐释《分级定义》的制定目的和标准。

转到粉色，最终变为橙红色。

团队对一些实际样品进行了测试，检测结果与实验室的检测结果基本一致。

目前，人们通常采用实验室仪器或酶抑制法等方法检测农残，但这种检测多由专业人员完成，检测仪器昂贵，检测结果要 2~3 天才能出来。

蒋长龙表示，其团队研发的传感器更加快速便捷，未经专业培训的人也可操作使用，可实现现场或实时检测，适用于基层环境监管部门、农贸市场及超市、个体消费者。

“该方法的创新之处在于结合智能手机对荧光信号进行处理，方便、准确。此外，该传感体系使用两种荧光物质作为信号，而不是依赖于酶，在现场检测中具有有一定应用潜力。”一位审稿人如是说。

不过，此次研发的传感器仅针对草甘膦残留检测。此外，传感器的检测信号依赖于宽光谱荧光色度的变化，而这种荧光色度可能会受到使用环境中光的影响。

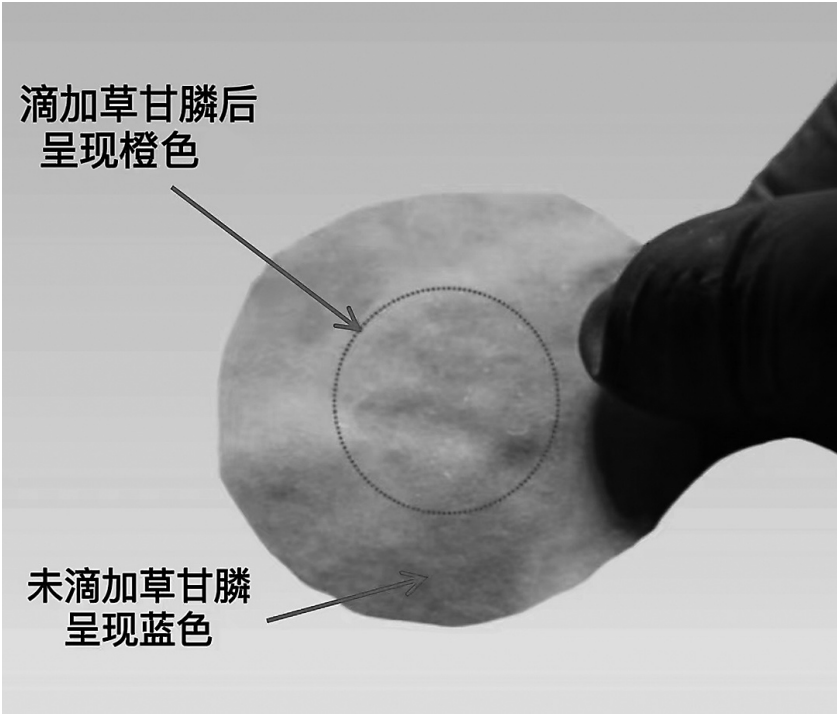
下一步，研究团队将探索制备多色发光纳米探针，进一步构建针对多种目标分析物的快速检测平台，研制纳米光效应传感器器件，用于制定环境中多种污染物检测的评价体系与技术标准，以期在人体健康预警可视化分析检测方面取得新进展。

相关论文信息：

<https://doi.org/10.1016/j.jhazmat.2022.129320>

■按图索技

便携平台 2 秒测出农药残留



本报讯(见习记者王敏)近日，中科院合肥物质科学研究院固体物理研究所研究员蒋长龙团队开发了一种新型无酶的便携式传感平台(以下简称传感平台)，2秒内检测出环境和食品中的草甘膦残留，最终浓度结果直接显示在智能手机上。相关研究成果发表于《危害物质杂志》。

“将瓜果蔬菜表面润湿，用检测试纸轻轻擦拭，约 2 秒后，用紫外灯照射，人们就可以通过试纸颜色变化大致判断草甘膦残留浓度的高低。”蒋长龙告诉《中国科学报》，如果试纸是蓝色，说明草甘膦残留浓度很低；如果试纸是粉色，说明浓度较高；当试纸是橙红色时，说明浓度很高。

“这种方法属于初筛，适合人们居家自测。”蒋长龙说，若想得到更精确结果，需要将试纸放入传感平台的试纸槽内。使用传感平台自带的紫外灯照射试纸，再用手机拍摄试纸照片，经手机的颜色识别软件自动分析转换后，最终显示农残浓度超标与否。

传感平台包括传感器、可用于读取数据的智能手机，提供荧光检测环境的手机附件。

“传感器是主要‘功臣’，由团队设计制备的蓝色碳点和金纳米团簇构成，能快速识别检测草甘膦。”蒋长龙说。其原理是草甘膦进入传感器后与碳点反应，导致碳点的蓝色荧光快速猝灭，而金纳米团簇的橙色荧光保持不变。从视觉上来看，试纸颜色从蓝色

课题组人员用试纸检测草甘膦。

课题组供图

■专家讲坛

中国人的高血压为何“与众不同”——中国脑卒中一级预防研究的启示

■霍勇

《中国心血管健康与疾病报告 2021》指出，目前我国心血管病患者高达 3.3 亿，其中高血压患者约 2.45 亿，如此庞大的高血压患病群体使我国脑卒中、急性心肌梗死等心脑血管急症的发病率居高不下，也给我国的慢病防治体系带来了巨大挑战。

我国高血压病呈显著的“三高三低”特点：“三高”即患病率高、增长趋势高、危害性高；“三低”即知晓率低、治疗率低、控制率低。

放眼全球，不少欧美国家的高血压患病率虽高于我国，但高血压和脑卒中的发病率却远低于我国。这究竟是因为我国高血压防治水平落后，还是我国人群高血压与欧美国家人群的高血压有所不同？带着这样的疑问，我们开启了长达 20 余年的探索。

首次提出“H 型高血压”

从 1995 年开始，我们研究团队就着手建立了涵盖 3.9 万人、跟踪随访时间长达 6.2 年的流行病学队列。我们通过大量数据分析发现，我国人群整体叶酸水平偏低，血同型半胱氨酸

(HCY)水平偏高，两者代谢通路中亚甲基四氢叶酸还原酶基因突变率高，三者与高血压联合作用，是我国脑卒中高发的重要原因。

2008 年，我们首次提出“H 型高血压”的概念，即“伴有血同型半胱氨酸升高的高血压”。这一概念的提出对心脑血管急、慢病防治具有非常重要的意义。

同年，我们开启了针对我国人群 H 型高血压高发特点设计的中国脑卒中一级预防(CSPPT)研究，这是迄今全世界最大的补充叶酸进行脑卒中一级预防的随机、双盲、对照临床研究。

CSPPT 研究历时近 5 年，共纳入 20702 例无脑卒中和心肌梗死病史的中国成年高血压患者。根据亚甲基四氢叶酸还原酶 C677T 基因型分层，我们将患者随机、双盲分成两组，分别每日服用单片固定复方制剂马来酸依那普利叶酸片(由 10mg 依那普利和 0.8mg 叶酸组成)和单纯依那普利(10mg)，主要观察终点是首发脑卒中，次要终点是首发缺血性脑卒中、首发出血性脑卒中、心肌梗死、全因死亡，以及包括心血管死亡、心肌梗死和脑卒中在内的心血管事件复

合终点。

CSPPT 研究结果证实，相较于服用依那普利的对照组，服用马来酸依那普利叶酸片的试验组在血压控制一样的情况下，额外降低首发脑卒中风险 21%。该研究 2015 年发表在《美国医学杂志》上。

“荣成模式”效果凸显

CSPPT 研究有着重要意义及成果转化效益。相关研究成果发表之后，很多国家及地区的研究都进一步证实，补充叶酸可以降低脑卒中的发病风险，从而也进一步证实 CSPPT 研究成果具有可重复性。

对于公众而言，CSPPT 研究传递的信息主要有三点。第一，我国人群的高血压有不同于欧美人群的显著特点。研究发现，中国人群体内叶酸含量低，叶酸处于低水平的人群占绝大部分，这也是我国高血压心脑血管病风险防治效果始终未达到理想状态的主要原因之一。第二，补充叶酸可有效减少脑卒中的发生。对于高血压患者尤其是 H 型高血压患者而言，CSPPT 研究给

■集装箱

3 种 4500 米级深海原位荧光传感器海试成功

本报讯(见习记者孙丹宁)近日，由中科院大连化学物理研究所研究员耿旭辉、关亚风团队与中科院深海科学与工程研究所相关团队共同研制的 4500 米级深海原位叶绿素荧光传感器、有色溶解有机物荧光传感器和微生物荧光传感器，近日在“探索二号”南海科考航次中搭载深海原位实验室进行了 24 次海底试验。其最大潜深 1833 米，并在海底连续工作 7 天，均获得了有效数据。

本次海试中，3 种传感器分别测量到了南海海水中从海平面到海底整个剖面的叶绿素 a、有色溶解有机物和微生物的浓度，包括剖面浓度变化趋势、拐点深度和绝对浓度。其测量结果均与文献报道数据相吻合，表明这 3 种传感器的测量及标定精度良好。其中，微生物传感器经第三方测试，对色氨酸检测下限为 0.03 μg/L，与国外同类高端产品水平相当。

青岛首个康复产业孵化器开园

本报讯(记者廖洋 通讯员肖强)日前，青岛市首个康复领域专业孵化器——青岛康复产业孵化器在青岛高新区正式开园。青岛市副市长耿涛，中国工程院院士、康复大学(筹)校长董尔丹出席并致辞。

该孵化器以康复大学(筹)建设为契机，在青岛市科技局支持下，由青岛高新区打造、青岛市工业技术研究院负责运营建设，致力于引进海内外高科技成果和高层次人才队伍，打造以康养为特色的生物医药及医疗器械产业集群，助力青岛高新区打造“中国康谷”。

据介绍，该孵化器计划在 3~5 年内建成全国知名的康复专业孵化器、科技成果转化示范区、康复产业科技创新高地，全面深化青岛高新区“中国康谷”品牌打造，将构建从科技成果转化直至企业上市的完整培育链条，为康复企业提供一站式、定制化加速服务，助力企业快速成长。

围绕康复产业生态搭建，由康复产业孵化器发起的青岛市康复产业创新发展联盟也正式宣告成立。联盟链接 30 余家海内外高校院所、领军企业、医疗机构、投资机构，将进一步整合产业创新要素、赋能产业发展。

《2021 年国外核工业与技术重大发展动向》发布

本报讯(记者张晴丹)日前，中国核学会、中核战略规划研究总院(中核智库)、中国原子能科学研究院等 3 家单位在北京举行了《2021 年国外核工业与技术重大发展动向》(以下简称《动向》)联合发布仪式。中核战略规划研究总院首席专家许春阳发布《动向》研究成果。

该研究成果聚焦核战略与核武器装备、核军控与防扩散、核能与核燃料、核技术应用、退役治理等五个领域，重点关注 2021 年国外发生的重大动向，本着“针对性”“重大性”“标志性”的原则，遴选出 40 多个具有现实意义或潜在影响的专题，通过撰写研究报告，反映国外核工业与技术领域战略规划、工业能力、装备技术等方面的重大发展动向，供有关部门、核工业从业人士和社会公众参阅。

《动向》指出，各国普遍将核能视为推动“双碳”目标的重要解决方案之一，在推进在运反应堆延寿、大型压水堆改进部署的同时，积极发展小堆、快堆、聚变堆、耐事故燃料等先进核能技术，提高核电安全性、经济性，确保核能可持续发展。

我国生物航煤实现首次规模化工业试生产

本报讯(记者计红梅)记者日前从中国石化新闻办获悉，中国首套生物航煤工业装置在中国石化镇海炼化首次产出生物航煤。这意味着我国生物航煤可实现规模化生产，向大规模生产及商业化应用迈出坚实一步。与传统石油基航空煤油相比，生物航煤全生命周期二氧化碳排放最高可减排 50%以上。该装置年设计加工能力 10 万吨，基本能消化一座千万人口城市一年回收的地沟油，每年可减排二氧化碳约 8 万吨，相当于近 5 万辆经济型轿车停开 1 年。

该套生物航煤工业装置采用中国石化石油化工科学研究院研发的生产技术。其使用用过的烹饪用油、食用油，即俗称的“地沟油”，作为原料生产生物航煤，“地沟油”含有大量脂肪族类化合物，含氧量高达 11%左右，氧分子会直接影响炼化装置催化剂的活性和稳定性，而传统的原料——石油含氧量低至 0.1%，相差超过 100 倍。此外，地沟油还含有硫、氯、金属元素等各类杂质，都要——去除。

为此，中国石化自主开发了专用催化剂和工艺，并经过了工业规模装置示范验证。今年 5 月，中国石化镇海炼化油脂加氢路线生物航煤产品通过可持续生物材料圆桌会议(RSB)认证。此次认证是我国生物航煤产品获得的第一张全球可持续性认证证书，表明该生物航煤装置原料、生产工艺及产品均符合 RSB 生物燃料可持续发展的基本原则与标准。