

## 人工智能专家： 未来须确保AI系统具人类意识且无偏见

叶伟强 报道  
yapwq@sph.com.sg

美国人工智能专家说，人工智能的发展将趋向与人类有更多互动交流，但这意味着有更多问题须防范和解决，包括如何确保这些新科技系统具备人类意识，且不带偏见。

于2016年至2018年担任国际人工智能协会会长的美国亚利桑那州立大学计算机科学教授苏巴拉奥（Subbarao Kambhampati），昨天受邀到新加坡管理大学开讲。

他说，一般人对人工智能（Artificial Intelligence，简称AI）

科技的心态是矛盾的，一方面想借助它提升工作生活的品质并从中受惠，另一方面却担心它取代工作等问题。

随着AI的普及化，AI系统必须做到能和人类用户“相辅相成”，因此下来须侧重于“人类意识”，也就是得加入人类情感和社会智能，包括“在特定情况时，给用户正确的情绪反应”。

“但其中一个得注意的是AI系统是根据用户输入的数据资料来学习，因此得确保数据是全面、而非偏向任何一方，以免系统带有偏见。”

他举例，有研究显示某个

自动驾驶车公司的系统无法很好地辨认皮肤偏暗褐色者，可能危及包括他自己的行人。

“此外，医疗护理数据可能以当地人数最多的种族为主，但这可能导致AI系统无法很好地照顾到其他族群。”

苏巴拉奥也说，尽管AI科技开创了许多可能，但它被攻击的可能也大幅增加。

“例如目前已有方式可骗过一些能辨认图像的AI系统，虽然用户一眼就看得出图像的物品，但系统反而被误导而认错物品。试想想，自动驾驶车如果把停车指示‘看成’加速指示，祸害会很大。另有AI系统可制造出神似真人脸孔的图像，这意味着人们可制造假身份。用假声音配真视频也是略有所闻的。”

他最后说道：“但我还是觉得AI有非常多广泛的用途，只是要当心。”



美国亚利桑那州立大学计算机科学教授苏巴拉奥（右）分享“研究有人类意识人工智能系统”的挑战，交流会由新加坡管理大学校长江莉莉教授主持。（陈斌勤摄）

其中一个得注意的是AI系统是根据用户输入的数据资料来学习，因此得确保数据是全面、而非偏向任何一方，以免系统带有偏见。

——美国亚利桑那州立大学  
计算机科学教授苏巴拉奥