

暴露AI风险 加剧人们对科技矛盾心理

尽管对于人工智能的担忧由来已久，但南洋理工大学研究人员通过“越狱”诱骗机器回答“不良”问题后，不少人加剧了他们对新科技的矛盾心理。

本地人机结合翻译公司译力亚洲（Elite Asia）首席执行官方莺吟接受《联合早报》访问时说，以前靠人工很难在一个晚上翻译100多页法律文件，但人工智能（AI）大大提高了工作效率。

“然而，南大研究所涉及的‘越狱’问题，可能会让机器产生错误甚至有害的信息，给企业造成损失。当一些企业利用免费机器服务，可能无意间将企业信息对外泄露，为

‘越狱’提供条件。”为了避免这些问题，她认为增设私人服务器和人工把关尤为重要。

在本地从事翻译校对工作八年的李维娜（35岁）告诉记者，机器相对于人工翻译具有不会遗漏、用词一致性等优点，但她也担心人工智能被用来造假甚至制造武器。

“最近在阿联酋举行的全球媒体大会期间，一些参展方展示了如何依靠一个人的照片和一段录音，就能让他用不同语言播报假新闻。”

新加坡教师陈炜杰（39岁）拥有16年教学经验。由于拥有国大计算机专业学位，他善于将数学模型和科技融入教学。虽然对“越狱”有所耳

闻，但南大的发现让他有些不安。

“南大这项技术展示了人工智能的非凡自我进化能力，但也暴露了人工智能被操纵以推翻伦理和操作准则的危险。”

他认为，人类必须对人工智能社区保持警惕，做到伦理前瞻并通过协作解决问题，确保人工智能的进步不会削弱社会和伦理框架。

谷歌前高管曾对AI风险发出警告

新加坡全国人工智能核心（AI Singapore）人工智能产业创新总监廖永健受访时说，谷歌前高管杰弗里·欣顿

（Geoffrey Hinton）和其他科学家曾对人工智能风险发出的警告言之有理。人工智能颠覆性的能力倘若使用不当，将沦为传播假新闻，发起网络攻击，甚至大量生成偏激言行的工具。

“我们必须记住，人工智能和其他强大的高科技一样，取决于如何正确使用。”

他告诉记者，AI Singapore致力于创立人工智能产品和解决方案，并利用这些项目对国人开展人工智能人才培养计划（AI Apprenticeship Programme）。这些项目经过严格审核，让受训者充分感知有关人工智能的伦理和安全问题。