

# 人工智能课程中伦理问题之考量



汇报人：华中师范大学 沈建铭

# outline

## 一 AI提出的一些伦理问题

- AI如何在我们的社会中行事？
- 如果工作供不应求，我们该怎么办？
- AI系统被允许死人吗？
- 我们应该关注超级智能和奇点吗？
- 我们应该如何对待AI？

## 二 关于伦理学和人工智能的工具

- 道义学 功利主义 德性伦理学

## 三 课堂伦理理论：充分发挥多角度

- 案例研究：老年护理机器人 弗兰克和机器人

## 四 AI课堂伦理

- 伦理道德问题 案例探究 教学资源



# AI提出的一些伦理问题

- AI如何在我们的社会中行事？
- 如果工作供不应求，我们该怎么办？
- AI系统被允许杀人吗？
- 我们应该关注超级智能和奇点吗？
- 我们应该如何对待AI？



# 关于伦理学和人工智能的工具

道义学  
功利主义  
德性伦理学

伦理学的三种方法

值得注意的是，这三种方法不一定是，也不应该被视为独立的或不排他的。我们不是争论任何特定系统的优越性；事实上，我们认为，彻底的道德教育将使學生能够利用所有三个主要的理论，在某些情况下将它们结合起来。



# 道义学

- 人的行为必须遵照某种道德原则或按照某种正当性去行为的道德理论。
- 侧重的是道德行为动机，不注重行为的后果，核心是义务和责任。
- 康德：“为义务而义务”，绝对律令



# 功利主义

- 功利主义认为人应该做出能“达到最大善”的行为，所谓最大善的计算则必须依靠此行为所涉及的每个个体之苦乐感觉的总和。
- 侧重结果，不关心行为与动机
- 边沁最大多数人的最大幸福



# 德性伦理学

- 关于人的品格的判断作为最基本的道德判断的理论。
- 关注行为主体，整体判断
- 亚里士多德，关注美德：智慧、勇毅、节制、正义



	侧重点	代表人物	代表观点
道义论	义务、责任	康德	要只按照你同时也能成为普遍规律的准则去行动
功利主义	结果	边沁	最大多数人的最大幸福
德性伦理学	行为者	亚里士多德	四大美德"智慧,勇敢,正义,节制





# 案例研究：老年护理机器人



老人弗兰克喜欢偷东西，但是他患了老年痴呆，记忆力慢慢衰退，儿子给他带来个机器人照顾他的生活。在与机器人的相处中，机器人为他做饭，收拾屋子等，更重要的他教会了机器人开锁，两人一起去偷东西。直到最后机器人与弗兰克一起去偷了很贵重的珠宝。

丢珠宝的人和警察找上门来，他支走了看望他的女儿，让机器人替他说谎叫来儿子，把警察耍了。最后他和机器人出去一圈后回到家里，警察被关在门外，机器人又再一次的让弗兰克把他格式化，清除他的记忆，让警察没有他们俩去偷东西的证据。最后在机器人的劝说和指导下机器人被格式化了。

弗兰克住进了养老院，儿子来看他，送儿子回来，看到一个机器人，他愣了一下，以为是他的那个，随后转角处又出现了一个一模一样的，他知道想多了，进屋了。完。



# AI课堂伦理

学生应该

- (1) 思考AI技术和系统提出的伦理问题，
- (2) 学习伦理学理论（道义学，功利主义和德性伦理学），提供框架，使他们能够思考伦理问题，
- (3) 将他们的知识应用于一个或多个案例研究，以描述其中发生的情况，并考虑可能解决他们所构成的伦理问题的方法



# AI课堂伦理

## 伦理道德问题

比如：

我们需要担心其可靠性，鲁棒性和安全性吗？

我们是否需要对其运作提供监督或监督？

我们如何保证他们的行为符合社会规范和人类价值观？

我们如何确定AI何时作出“错误”的决定？  
谁对这个决定负责？

我们应该如何测试他们？

我们应该使用哪些应用程序？

谁在生活水平，分配和工作质量以及其他社会经济因素方面受益匪浅？



# 结论

我们的立场是，我们作为教育工作者有责任培养学生的认识到他们作为技术专家可能遇到的更大的伦理问题和责任，以科幻作为基础为基础，实现了更好的学生学习，保留和理解。



# NARS的伦理设计

- NARS教育
- 先天设计

