人工智能与安全

课程背景：

2017 年 3 月，人工智能首次被写入《政府工作报告》，同年 7 月，国务院颁发《新一代人工智能发展规划》，提出了“三步走”的战略目标，宣布举全国之力在 2030 年抢占人工智能全球制高点。今天的中国社会，智能应用已经随处可见，我们正处在智能化的社会之中。

当今人工智能科学更准确的是指数据智能，在大数据时代人工智能技术应用得到了非速发展，无论是计算智能，还是感知智能，都已为产业界各类创新提供主要技术支撑，甚至出现认知智能技术的初级尝试应用。所以可以断定，在当今的中国产业界，不了解大数据与人工智能的科技应用，大概率会在竞争中处于劣势。

但是，我们也常常听身边的企业家员表示出对智能类应用的安全性担忧，比较有代表性的埃隆·马斯克（Elon Musk）、史蒂芬·霍金（Stephen Hawking）和比尔·盖茨（Bill Gates）都忌惮人工智能（以下简称“AI”），并忧心忡忡地表示在不久的将来，AI 会对人类的生存构成威胁。埃隆·马斯克曾宣称人工智能是人类文明持续存在所面临的重大危机。而为什么另一个同样杰出的群体，包括马克·扎克伯格（Mark Zuckerberg）、吴恩达（Andrew Ng）和佩德罗·多明戈斯（Pedro Domingos）在内，又认为“人工智能威胁论”给的理由太过牵强，几乎不攻自破。扎克伯格甚至称那些鼓吹末日将至的人相当不负责，当今人工智能领域最伟大的人物之一吴恩达亦表示，这种焦虑就像是担忧“火星上的人口过剩”，完全没有必要……

本课程开课的主要目的是面向非专业人士，尤其是产业界，从人工智能产生的背景、基本原理、技术体系入手，摒弃社会上流行的各类带有商业引导目的的内容，通过大量丰富案例作证，系统性讲授人工智能及人工智能安全问题，对非专业技术人员，能够听得明、学得会、用得好。在深度推进产业智能化与企业智能化应用的同时，对人工智能技术应用带来的一系列安全问题有更为全面的认知与理解。本课程作为提升企业家数字化生存能力素质的重要组成部分。

课程收益：

- 厘清大数据、人工智能体系中的关键问题；
- 熟悉人工智能行为体的分类及对应安全问题；
- 了解世界各国人工智能发展战略规划与安全伦理准则；
- 熟悉社会常见的人工智能应用，强化对智能的认知；
- 熟悉掌握新兴技术成熟度曲线，把握未来技术发展趋势；
- 了解人工智能主要流派与基本原理、主流神经网络模型；
- 熟悉人工智能安体系与架构及通用人工智能的发展趋势。

课程时间：2 天，6 小时/天
课程对象：
● 创业者、企业负责人、企业创始团队、董事会成员
● 企业高级管理人员，总经理、总工程师、副总经理等；
● 渴望掌握新兴技术价值落地的企业中层以上管理人员及工程师；
● 高校 MBA、EMBA、DBA 专业研究生；
● 地方政府中、高级管理干部及相关领域公务员；
● 国家“十四五”规划中智能制造及战略新兴产业中相关人士。
课程方式：理论+案例+实操+演练
课程风格：
源于实战：以客户需求驱动的咨询引导型培训，以最前沿科技和典型案例演练启迪学员；
逻辑性强：理论、实践、研究成果高度结合，用通俗易懂的语言使各类学员听懂并掌握；
深入浅出：现场教学既幽默风趣又富有哲理，结合研究成果和实践经验进行现身说法；
价值度高：课程内容经过市场实战打磨，是学员由外行变成内行的知识利器；
方法论新：经过专门面向非专业人士设计,专业知识+刻意练习+行动学习+问题改善工作坊,对不同学员的诉求一律耐心互动，并能够为大客户实现授课与顾问、工程服务相结合。

课程大纲

导入 1：现在是大数据时代，现在是人工智能时代
案例：“我的一天”
研讨：(GP-分组对抗记分点)感受智能化，分组讨论描述“你的一天”，并指出哪些应用或名词是和人工智能紧密相关？（除老师事先讲过的，答对一条记 1 分）
导入 2：“人工智能威胁论”
案例：人工智能的技术发展，对人类来说到底是生存还是毁灭？
研讨：(GP-分组对抗记分点)分组汇总每位同学的观点，提交给老师。
备注：GP-为短时时间讨论，一般不超过 5 分钟，LGP 为长时间讨论，一般在 5-20 分钟；GP 活动由老师根据现场情况发起或不发起，非固定活动。下同。

第一讲：大数据时代特征与人工智能国家战略

一、从互联网到大数据时代的演变过程

1. 从互联网、Web2.0、移动互联网看人类在线化过程
2. 人类在线化过程与行为数据的关系

案例分析：以商业购物场景为例，分析人类活动的在线变化及其产生的行为数据
小组研讨：(GP)分组设计其他场景,延伸到物联传感网，并总结，老师点评并打分

3. 大数据的来源与全球数增长情况分析
4. 数据计量单位的换算
5. 5G 的战略地位与价值
6. 大数据的两个重要特征
7. 大数据价值的现状

二、国际与中国人工智能发展

1. 中国：人工智能的国家战略与“智能+”
2. 世界各国人工智能发展对比分析
3. 解读“十四五”规划给我们的启示

小组研讨：(LGP)找出所在行业的有关人工智能方面的国家或地方政策规划，分析原因与机会

第二讲：人工智能发展史

一、人工智能的起源
1. 人工智能产生的背景
2. 图灵与图灵测试
3. 达特茅斯会议与“人工智能”

二、人工智能的三次浪潮
1. 第一次人工智能浪潮：推理与探索

案例分析：计算机在使用“推理和探索”的兴起与没落

2. 第二次人工智能浪潮：知识工程

案例分析：专家系统的窘境与问题

3. 我们正在第三次人工智能浪尖上：大数据与深度学习

案例分析：人工智能发展历程中的里程碑事件

第三讲：人工智能原理

一、人工智能定义与分类
1. 人工智能的定义与正确理解
2. 计算智能、感知智能与认知智能
3. 人工智能的几大门派其及技术发展方向

二、人工智能人才与知识体系
1. 学科领域交叉与渗透下的人工智能创新协同
2. 世界及中国人工智能类人才培养现状

案例分析：中国某顶尖大学人工智能研究院体系及研究领域

3. 把握与跟踪人工智能技术发展趋势的方法

案例分析：深度分析 Gartner 曲线

实操演练：(LGP)依据现场给出的某人工智能应用,依据 Gartner 曲线分析其技术发展规律与特点

三、数据智能平台技术体系
1. 大数据技术平台架构
2. 人工智能技术平台架构
3. 通用深度学习开源框架与特点

第四讲：常见深度学习模型与应用

一、传统数据模型与应用
1. 常见传统数据算法与模型
2. 常见传统数据算法的应用

二、深度神经网络(DNN)模型与应用
1. DNN 模型
2. DNN 应用场景：搜索排序、推荐排序

三、卷积神经网络(CNN)模型与应用
1. CNN 模型
2. CNN 应用场景：图像识别、视频分析

四、循环神经网络(RNN)模型与应用
1. RNN 模型
2. RNN 应用场景：语音识别、自然语言处理

案例分析：人机智力大战的巅峰——阿尔法狗

第五讲：机器人技术及其应用原理（选讲课程）

一、机器人概述
1. “robot”一词的来源
2. 机器人定义与相关概念
3. 机器人发展历程
4. 机器人分类

二、机器人基本原理及应用
1. 机器人控制系统的基本结构
2. 工业机器人
3. 农业机器人
4. 医疗机器人
5. 服务机器人
6. 特种机器人

案例分析：机器人在工业、农业、医疗等领域的应用

第六讲：人工智能与安全哲学
一、安全
1. 人类社会对安全的认知与理解
2. 人工智能安全、人工智能与安全

二、通用人工智能安全的哲学命题
1. 技术革命视角下的人类四次纪元
2. 第四次纪元的不可控因素
3. 人工智能的安全命题
4．通用人工智能的三个哲学命题

小组研讨：（LGP）基于哲学上的命题，分组讨论形成各自主张。

第七讲：人工智能产业生态与安全
一、人工智能产业生态
1. 人工智能应用领域
2. AI 芯片与视觉传感器
3. AI 通用技术

案例分析：主流机器视觉、语音识别、自然语言、知识图谱应用的市场与趋势

二、狭义人工智能安全
1. 人工智能的安全体系
2. 人工智能的安全伦理概要
3. 人工智能安全对社会的冲击
4. 人工智能安全 VS 网络安全 VS 信息安全

案例分析 1：全球首例自动驾驶车辆撞死行人的案件
案例分析 2：《2020 年度全球十大人工智能治理事件》

第八讲：人工智能内生安全
一、数据安全
1. 数据投毒与反制
2. 对抗样本攻击与反制措施
3. 数据质量与数据安全之间的管理问题
4. 对产业界及管理者的启示

案例分析 1：深网视界曝出数据泄露事件
案例分析 2：地下产业链之数据隐私市场

案例分析 3：网上热传的几家著名科技公司的安全事件解读

二、算法与模式安全

1. 算法的可解释性与安全
2. 模型存储与管理的安全问题
3. 开源模型的安全问题
4. 对产业界及管理者的启示

案例分析 1：一支激光笔是如何打败了自动驾驶？

案例分析 2：医疗领域人工智能诊断技术应用的尴尬

三、框架与运行安全

1. 架框安全问题
2. 主观与客观原因上的运行安全与保障问题
3. 对产业界及管理者的启示

案例分析：几起自动驾驶车祸背后的安全分析

小组研讨：（LGP）结合分组学员企业的情景，研讨应用人工智能内生安全的思路。

第九讲：人工智能衍生安全与伦理

一、人工智能衍生安全

1. 智能系统失误引发的安全事故

案例分析 1：当自动驾驶、智能机器人、智能音箱、医疗机器人失效后……

案例分析 2：聊天机器人的偏激言论引发的群体影响

2. 人工智能行为体失控要素分析

案例分析：“机器人三定律”

3. 国际上预防人工智能技术失控的举措

二、人工智能伦理

1. 人工智能体是否应该赋予“人权”？
2. 通过使用人工智能的人权侵犯问题
3. 人工智能是否能成为伦理主体
4. 人工智能的伦理责任问题

案例分析：几起自动驾驶案件的责任追究

小组研讨：(GP)现实生活或科幻电影中的“智能人”，及其引发伦理的故事关键词。

三、人工智能伦理准则

1. 世界各国关于人工智能技术发展的伦理准则
2. 人工智能技术伦理准则的共识性与争议性
3. 我国专门提出人工智能伦理与法律的“三步走”规划

案例分析：解读《2020 年度全球十大人工智能治理事件》的处理结果

结束语：人工智能安全的未来展望!