

# 中国社会科学院大学“人工智能伦理与法治导论”课程大纲

课程基本信息 (Course Information)					
课程编号 (Course ID)	102042021083	*学时 (Credit Hours)	32	*学分 (Credits)	2
*课程名称 (Course Name)	(中文) 人工智能伦理与法治导论				
	(英文) Introduction to AI Ethics and the Rule of Law				
先修课程 (Prerequisite Courses)					
*课程简介 (Description)	<p>本课程坚持以习近平法治思想和习近平总书记关于网络强国重要思想为指导,落实立德树人的使命,让学生掌握人工智能伦理和法律的基本知识,提升学生运用法治思维思考和研究人工智能技术快速发展引发的各类问题,为未来从事相关领域的工作提供坚实的知识基础。</p> <p>本课程的目标包括:掌握人工智能伦理和法律的基本概念和原则;通过分析国内外政策和法律,使学生了解中国在人工智能领域的国家战略和政策导向,在人工智能治理伦理和法治领域贯彻人类命运共同体理念;引导学生关注人工智能技术对环境和社会的影响,激发学生的创新思维,探索绿色人工智能的发展路径;引导学生树立科技创新与社会责任相统一的价值观念,推动人工智能技术的良性发展;通过促使学生主动参与课程活动,加深对人工智能伦理与法治问题的理解和思考,培养学生的独立思考、团队合作和批判性思维的能力。</p> <p>本课程的主要内容包含:人工智能概述与发展历程、人工智能发展战略、人工智能伦理问题、人工智能的法律影响、数据伦理与法治、算法伦理与法治、算力治理与法治、域外人工智能治理以及人工智能的应用场景与法律治理等。本课程的重难点在于了解人工智能的技术特征的前提下,掌握当前人工智能发展战略和伦理原则,并能结合实例提出自己对于技术与法律之间互动的思考。</p> <p>本课程的基本要求:达到课程考勤指标,不得无故缺课、迟到、早退;积极参与课堂讨论,展开对人工智能伦理与法治问题的思考和辩论;对人工智能应用中的具体伦理与法治问题进行分析;完成课程论文,对一个特定的人工智能领域的伦理或法治问题进行深入研究和探讨。</p>				
*课程简介 (Description)	<p>This course is guided by Xi Jinping Thought on the Rule of Law and General Secretary Xi Jinping's Important Thinking on building the nation into a strong cyberpower, and aims to fulfill the mission of cultivating students' moral character and enabling them to master the basic knowledge of artificial intelligence ethics and law. It aims to enhance students' ability to think and study various issues caused by the rapid development of artificial intelligence technology from the perspective of rule of law, and provide a solid knowledge foundation for future work in related fields.</p> <p>The objectives of this course include: to master the basic concepts and principles of artificial intelligence ethics and law; to understand China's national strategy and policy orientation in the field of artificial intelligence through analyzing domestic and foreign policies and laws, and to implement the concept of a community with a shared future for mankind in the governance ethics and legal domain of artificial intelligence; to guide</p>				

	<p>students to pay attention to the impact of artificial intelligence technology on the environment and society, stimulate students' innovative thinking, and explore the development path of green artificial intelligence; to guide students to establish the value concept of the unity of technological innovation and social responsibility, and promote the positive development of artificial intelligence technology; to deepen students' understanding and thinking on artificial intelligence ethics and legal issues by encouraging them to actively participate in course activities, and to cultivate students' independent thinking, teamwork, and critical thinking abilities.</p> <p>The main content of this course includes: an overview and development history of artificial intelligence, development strategies for artificial intelligence, ethical issues in artificial intelligence, legal impacts of artificial intelligence, data ethics and legal governance, algorithm ethics and legal governance, computing power governance and legal governance, overseas artificial intelligence governance, and applications scenarios and legal governance of artificial intelligence. The difficulties and key points of this course lie in understanding the technical characteristics of artificial intelligence, mastering current development strategies and ethical principles of artificial intelligence, and combining examples to propose one's own thoughts on the interaction between technology and law. The basic requirements of this course are: to achieve attendance indicators and not to be absent from class without reason, late, or leave early; to actively participate in classroom discussions, think and debate ethical and legal issues related to artificial intelligence; to analyze specific ethical and legal issues in the application of artificial intelligence; to complete a course paper that deeply researches and explores ethical or legal issues in a specific field of artificial intelligence.</p>		
<p>*教材 (Textbooks)</p>	<p>周辉等：《人工智能治理：场景、原则与规则》，中国社会科学出版社 2021 年版</p>		
<p>参考资料 (Other References)</p>			
<p>*课程类别 (Course Category)</p>	<p><input type="checkbox"/>公共基础课/全校公共必修课    <input type="checkbox"/>通识教育课    <input type="checkbox"/>专业基础课  <input type="checkbox"/>专业核心课/专业必修课    <input checked="" type="checkbox"/>专业拓展课/专业选修课    <input type="checkbox"/>其他_____</p>		
<p>*授课对象 (Target Students)</p>	<p>全校本科生</p>	<p>*授课模式 (Mode of Instruction)</p>	<p><input type="checkbox"/>线上，教学平台_____</p> <p><input checked="" type="checkbox"/>线下    <input type="checkbox"/>混合式    <input type="checkbox"/>其他</p> <p><input type="checkbox"/>实践类（70%以上学时深入基层）_</p>
<p>*开课院系 (School)</p>	<p>法学院</p>	<p>*授课语言 (Language of Instruction)</p>	<p><input checked="" type="checkbox"/>中文    <input type="checkbox"/>全外语_____</p> <p><input type="checkbox"/>双语：中文+_____（外语讲授不低于 50%）</p>

*授课教师信息 (Teacher Information)	课程负责人 姓名及简介		周辉, 中国社会科学院法学研究所网络与信息法研究室副主任(主持工作)、副研究员, 中国社会科学院大学法学院副教授、硕士生导师, 美国耶鲁大学法学院访问学者。国家行政学院博士后(公共管理专业、行政法方向)、北京大学法学博士(民商法专业, 网络法方向), 中国法学会网络与信息法学研究会常务副秘书长, 中国互联网协会个人信息保护工作委员会副秘书长, 参与《个人信息保护法》、《电子商务法》等法律起草论证, 主持起草《网信系统法治宣传教育第八个五年规划(2021—2025年)(专家建议稿)》《人工智能示范法(专家建议稿)》, 系交通运输部首批公职律师, 入选北京市法学会“百名法学英才”队伍建设计划。			
	团队成员 姓名及简介		刘灿华, 中国社会科学院法学研究所网络与信息法研究室助理研究员, 中国社科院文化法制研究中心副秘书长。 徐玖玖, 中国社会科学院法学研究所网络与信息法研究室助理研究员。			
学习目标 (Learning Outcomes)	<p>1. 学生应该能够理解人工智能伦理和法治的基本概念, 包括伦理价值观、法律法规、人权保护等内容, 并能够运用这些概念来分析和评估人工智能技术的应用。</p> <p>2. 学生应该能够了解人工智能技术在社会中可能引发的伦理和法治问题, 如隐私保护、数据安全、智能算法的公平性等, 并能够分析和讨论这些问题的根源、影响和解决方案。</p> <p>3. 学生应该具备科学道德意识, 明白科技发展与社会发展之间的关系, 以及自身在科技发展中的责任。同时, 学生应该有高度的社会责任感, 能够思考人工智能技术对社会、环境和个体的影响, 并探索可持续发展的路径。</p> <p>4. 学生应该能够运用批判性思维分析人工智能伦理和法治问题, 形成独立的观点和见解, 并具备解决实际问题的能力。通过案例分析、讨论和实践项目, 培养学生的分析、判断和解决问题的能力。</p>					
*考核方式 (Grading)	平时成绩: 占 30% 期末成绩: 占 70%, 以期末论文方式结果。					
<b>*课程教学计划 (Teaching Plan) 填写规范化要求见附件</b>						
周次	周学时	其中				教学内容摘要 (必含章节名称、讲述的内容提要、实验的名称、教学方法、课堂讨论的题目、阅读文献参考书目及作业等)
		讲授	实验课	习题课	课程讨论	
第一周	2	2				<p><b>第一章 概论</b></p> <p>第一节 人工智能的概念和范围</p> <p>一、人工智能的概念定义</p> <p>二、人工智能的发展和演变</p> <p>第二节 人工智能的技术</p> <p>一、机器学习与深度学习</p> <p>二、计算机视觉</p> <p>三、自然语言处理</p> <p>四、机器人自动化</p> <p>第三节 人工智能伦理的主要命题</p> <p>一、数据隐私和安全</p> <p>二、公平和歧视</p> <p>三、透明度和解释性</p>

					<p>四、责任与问责</p> <p>第四节 人工智能法治的框架</p> <p>一、数据</p> <p>二、算法</p> <p>三、算力</p> <p>四、责任</p> <p>具体内容：本次课程主要讲述人工智能的概念以及定义；人工智能的发展和演变；人工智能的技术包括机器学习与深度学习、计算机视觉、自然语言处理、机器人自动化等；人工智能伦理的范畴、人工智能法治的框架</p> <p>教学方法：讲授法</p> <p>阅读文献：</p> <p>(1) 国家新一代人工智能治理专业委员会：《新一代人工智能伦理规范》，2021年</p> <p>(2) 联合国教科文组织：《人工智能伦理问题建议书》，2021年</p>
第二周	2	2			<p style="text-align: center;"><b>第二章 人工智能发展战略</b></p> <p>第一节 中国的人工智能发展战略</p> <p>一、发展背景和动机</p> <p>二、创新驱动和核心技术</p> <p>三、产业发展和应用场景</p> <p>四、政策支持和投资促进</p> <p>第二节 美国的人工智能发展战略</p> <p>一、政府支持和投资计划</p> <p>二、创新生态系统和科技企业</p> <p>三、人才培养和教育体系</p> <p>四、核心技术研究与发展</p> <p>五、国际合作和竞争优势</p> <p>第三节 欧盟的人工智能发展战略</p> <p>一、法规框架和伦理准则</p> <p>二、创新生态系统和科技企业</p> <p>三、数据驱动和隐私保护</p> <p>四、人才培养和教育体系</p> <p>五、安全防范和风险评估</p> <p>具体内容：当前人工智能的发展呈现全球竞争的格局，面临着数据隐私与安全、伦理和法律规范、技术与社会影响等多重挑战。同时，跨学科融合和国际合作的机遇也提供了推动人工智能发展的重要动力。通过讲授中国、美国和欧盟的人工智能发展战略，使学生掌握当前人工智能的发展格局和挑战。</p> <p>教学方法：讲授法</p>

					<p>阅读文献：</p> <p>[1] 国务院：《新一代人工智能发展规划》，2017年</p> <p>[2] White House. Executive Order on Maintaining American Leadership in Artificial Intelligence. 2019</p>
第三周	2	2			<p style="text-align: center;"><b>第三章 人工智能的伦理原则</b></p> <p>第一节 联合国</p> <p>一、《人工智能伦理建议书》的诞生背景及意义</p> <p>二、十项主要原则</p> <p>第二节 中国</p> <p>一、中国人工智能伦理原则的发展历程以及现状</p> <p>二、中国人工智能伦理原则的主要内容</p> <p>第三节 美国</p> <p>一、《美国人工智能权利法案蓝图》的主要内容</p> <p>二、美国人工智能伦理原则的主要内容</p> <p>具体内容：讲授联合国《人工智能伦理建议书》、《美国人工智能权利法案蓝图》以及中国的相关文件中确立的人工智能伦理原则，引导学生讨论公平非歧视、透明度与可解释性原则的含义和价值</p> <p>教学方法：讲授法</p> <p>讨论题目：透明度与可解释性的异同</p> <p>阅读文献：</p> <p>（1）国家新一代人工智能治理专业委员会：《新一代人工智能伦理规范》，2021年</p> <p>（2）联合国教科文组织：《人工智能伦理问题建议书》，2021年</p> <p>（3）Luciano Floridi: The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities, Oxford University Press, 2023.</p>
第四周	2	2			<p style="text-align: center;"><b>第四章 人工智能的法律影响（上）</b></p> <p>第一节 对主体的影响</p> <p>一、法律服务和司法改革</p> <p>二、法律职业与就业前景</p> <p>第二节 对行为的影响</p> <p>一、法律判断与决策</p> <p>二、自动化合同和智能合同</p> <p>三、司法预测与裁判结果</p> <p>四、法律研究与信息检索</p> <p>具体内容：讲授人工智能对主体和行为的影响。人工智能对主体和行为的影响有以下几个方面：自动化与替代；责任与追责；司法预测与裁判结果；法律研究与信息检索。引导学生思考，法律行为的方式和效率随着人工智能技术的应用会发生什么变化，以及如何对其法律影响进行深入研究和探讨，以确保法律体系的适应性和公正性。</p>

						<p>教学方法：讲授法</p> <p>阅读文献： 周辉：《深化法院信息化建设，实现智能互通的公平正义》，《人民法院报》2022年7月21日第05版</p>
第五周	2	2				<p style="text-align: center;"><b>第四章 人工智能的法律影响（下）</b></p> <p>第三节 对客体的影响</p> <p>一、AIGC</p> <p>二、AI 作为辅助物</p> <p>第四节 内容</p> <p>一、权利和权力</p> <p>二、义务</p> <p>三、责任</p> <p>具体内容：讲授人工智能对客体的影响。人工智能的发展给法律领域带来了诸多挑战和调整需求。引导学生在隐私保护、透明解释、责任与法律责任、知识产权以及就业与劳动法等方面讨论如何合理制定法律规范和确定伦理原则，以推进人工智能的健康发展并保护个体和社会的权益。</p> <p>教学方法：讲授法</p> <p>阅读文献： 周辉：《算法权力及其规制》，《法制与社会发展》2019年第6期</p>
第六周	2	2				<p style="text-align: center;"><b>第五章 数据伦理与法治</b></p> <p>第一节 数据与人工智能</p> <p>一、数据之于 AI 的价值</p> <p>二、《人工智能研发运营体系（MLOps）实践指南（2023年）》</p> <p>第二节 数据质量</p> <p>一、数据规模和代表性问题-数据内嵌的价值观</p> <p>二、数据垄断与“数据孤岛”-整合障碍</p> <p>三、“脏数据”：重复、缺失、造假、失效等</p> <p>四、数据模型的偏差</p> <p>具体内容：本次课程讲授数据对于人工智能的价值以及数据对于人工智能安全的影响、高质量数据的稀缺性。指导学生掌握数据隐私和保护、数据质量和准确性、公平和非歧视性、透明度和知情同意、跨境数据流和适用法律以及数据使用的社会影响等内容。</p> <p>教学方法：讲授法</p> <p>阅读文献： 中国信通院.人工智能研发运营体系（MLOps）实践指南（2023年）.http://www.caict.ac.cn/kxyj/qwfb/ztbg/202303/t20230316_416827.htm</p>

第七周	2	2				<p style="text-align: center;"><b>第六章 算法伦理与法治（上）</b></p> <p>第一节 算法与人工智能  一、算法原理与机制  二、算法的特征及规制难点</p> <p>第二节 算法的主要类型  一、算法的类型：排序、搜索、回溯  二、算法的类型：近似、优化、分治</p> <p>主要内容：本次课程讲授算法的原理和机制以及法律视角下算法不同类型的区别。通过介绍算法的含义以及在人工智能领域的应用，帮助学生快速掌握算法于人工智能的功能，为接下来的课程奠定基础。</p> <p>教学方法：讲授法</p> <p>阅读文献：  周辉：《算法权力及其规制》，《法制与社会发展》2019年第6期</p>
第八周	2	2				<p style="text-align: center;"><b>第六章 算法伦理与法治（下）</b></p> <p>第三节 算法的治理手段  一、算法公开  二、算法解释  三、增强个人自主权  四、技术治理</p> <p>具体内容：主要围绕公平性和歧视、透明性和可解释性、隐私保护和数据安全、责任与法律责任以及社会影响和公共参与等内容展开，使学生理解应当确保算法的设计、开发和应用符合伦理原则和法律规定，最大程度地保护个体和社会的权益。</p> <p>教学方法：讲授法</p> <p>阅读文献：  周辉：《算法权力及其规制》，《法制与社会发展》2019年第6期</p>
第九周	2	2				<p style="text-align: center;"><b>第七章 算力治理与法治</b></p> <p>第一节 算力与人工智能  第二节 算力基础设施建设/关键基础设施的供应安全  一、数据中心与算力网络建设  二、芯片与服务器等硬件保障  三、云服务  第三节 绿色低碳的算力  一、绿色低碳的意义和特征  二、国内外的实践案例</p> <p>具体内容：本课程将引导学生关注人工智能技术对环境和社会的影响，激发学生的创新思维，探索绿色人工智能的发展路径，并通过案例分析和实践项目，帮助学生形成对可持续发展的责任意识。</p> <p>教学方法：讲授法</p>

					<p>阅读文献：</p> <p>(1) 工业和信息化部：《新型数据中心发展三年行动计划（2021—2023年）》</p> <p>(2) 国际数据公司(IDC)、浪潮信息：《2023-2024 中国人工智能计算力发展评估报告》</p>
第十周	2	2			<p style="text-align: center;"><b>第八章 人工智能的应用场景及其法律治理（上）</b></p> <p>第一节 智慧法治</p> <p>一、智慧立法</p> <p>二、智慧司法</p> <p>三、智慧执法</p> <p>第二节 自动驾驶</p> <p>一、技术发展和应用现状</p> <p>二、主要问题</p> <p>三、治理规范</p> <p>具体内容：本次课程讲授在智慧法治和自动驾驶应用场景下人工智能的治理问题。其中，智慧法治包括立法、司法和执法方面的回应；自动驾驶包括交通事故责任承担以及隐私和数据安全方面的问题。</p> <p>教学方法：讲授法</p> <p>阅读文献：</p> <p>(1) 郑志峰：《自动驾驶汽车的私法挑战与应对挑战》，中国法制出版社 2022 年版</p> <p>(2) 周汉华、刘灿华：《社会治理智能化的法治路径》，《法学杂志》，2020 年第 9 期</p>
第十一周	2	2			<p style="text-align: center;"><b>第八章 人工智能的应用场景及其法律治理（下）</b></p> <p>第三节 电子商务</p> <p>一、智能算法的应用</p> <p>二、突出问题</p> <p>三、监管重点</p> <p>第四节 生成式人工智能</p> <p>一、技术发展与应用现状</p> <p>二、挑战与风险</p> <p>三、监管案例与治理重点</p> <p>具体内容：通过讲授电子商务领域人工智能应用的特征和治理规范，引导学生探讨不同应用场景下人工智能治理的差异和共性，并结合生成式人工智能的应用提出自己的思考。</p> <p>阅读文献：</p> <p>(1) 世界互联网大会人工智能工作组：《发展负责任生成式人工智能研究报告及共识文件》，2023 年 11 月。</p>

第十二周	2	2				<p style="text-align: center;"><b>第九章：域外人工智能的法律治理——美国（上）</b></p> <p>第一节 美国人工智能治理的整体框架</p> <p>一、联邦重点立法</p> <p>二、州重点立法</p> <p>第二节 美国基础模型的新近发展：以 CHATGPT 为例</p> <p>一、Openai 与 CHATGPT</p> <p>二、Meta 与 Llama</p> <p>具体内容：介绍美国的人工智能治理情况，帮助学生理解美国人工智能治理政策和法律诞生的背景以及实施效果，引导学生思考美国人工智能治理对我国的启示。</p> <p>教学方法：讲授法</p> <p>阅读文献：</p> <p>《专访中国社科院法学所周辉：美国最新 AI 行政命令具有较强的前瞻性，未来可能出现“华盛顿效应”》，《21 世纪经济报道》2023 年 12 月 1 日</p> <p>《专访耶鲁大学 Karman Lucero：美国监管趋势收紧，人工智能国际治理共识讨论需保证透明》，《21 世纪经济报道》2023 年 12 月 4 日</p>
第十三周	2	2				<p style="text-align: center;"><b>第九章：域外人工智能的法律治理——美国（下）</b></p> <p>第三节 美国人工智能的监管政策</p> <p>一、美国人工智能行政命令</p> <p>二、美国人工智能风险管理框架</p> <p>第四节 小结</p> <p>一、美国人工智能治理的重点行业自律模式</p> <p>二、启示与借鉴</p> <p>具体内容：介绍美国的人工智能治理情况，重点介绍 Chatgpt 的治理以及美国在人工智能治理领域达成的共识，帮助学生掌握美国人工智能治理框架，并引导学生思考其优缺点以及对我国的启示。</p> <p>教学方法：讲授法</p> <p>阅读文献：</p> <p>（1）美国政府问责局：《人工智能：联邦机构和其他实体的问责框架》，2021</p> <p>（2）美国国家标准与技术研究院 (NIST):《人工智能风险管理框架》，2023</p> <p>（3）《美国关于确保人工智能安全、可靠和可信的行政命令》，2023</p>
第十四周	2	2				<p style="text-align: center;"><b>第十章 域外人工智能的法律治理——欧洲（上）</b></p> <p>第一节 《欧盟人工智能法案》</p> <p>一、立法背景与目标</p> <p>二、主要法律制度</p> <p>第二节 《欧盟人工智能责任指令草案》</p>

					<p>一、立法背景与目标</p> <p>二、主要法律制度</p> <p>具体内容：通过介绍《欧盟人工智能法案》、《欧盟人工智能责任指令草案》立法背景、目标和主要法律制度，指导学生把握欧盟人工智能治理概况，使学生理解人工智能应用中确保公共利益的保护、个体权益的尊重，并促进技术的负责和可持续发展。</p> <p>教学方法：讲授法</p> <p>阅读文献：</p> <p>(1) 《欧盟人工智能法草案》，2023</p> <p>(2) 《欧盟人工智能责任指令草案》，2022</p>
第十五周	2	2			<p style="text-align: center;"><b>第十章 域外人工智能的法律治理——欧洲（下）</b></p> <p>第一节 欧盟人工智能治理的特点与借鉴</p> <p>一、欧盟人工智能治理的特点</p> <p>二、欧盟人工智能治理的启示与借鉴</p> <p>二、欧盟人工智能治理的主要经验</p> <p>第二节 英国人工智能治理的特点</p> <p>具体内容：在此前课程基础上，引导学生总结归纳出欧盟人工智能的整体框架以及主要特点，并带领学生探讨欧盟和英国人工智能治理的特点和经验。</p> <p>教学方法：讲授法</p> <p>阅读文献：</p> <p>(1) 《欧盟人工智能法草案》，2023</p> <p>(2) 《欧盟人工智能责任指令草案》，2022</p> <p>(3) 周辉、金僖艾：1.《英国人工智能监管实践、创新与借鉴》，《数字法治》，2023年第5期</p>
第十六周	2	2			<p style="text-align: center;"><b>第十一章 基础模型企业风险治理</b></p> <p>第一节 国内基础模型企业风险治理</p> <p>一、百度与文心一言</p> <p>二、阿里巴巴与通义千问</p> <p>第二节 国外基础模型企业风险治理</p> <p>一、谷歌</p> <p>二、Openai</p> <p>三、Meta</p> <p>四、Microsoft</p> <p>第三节 基础模型企业自律治理的外部约束和社会监督</p> <p>一、外部约束</p> <p>二、社会监督</p> <p>具体内容：讲授国内外代表性基础模型企业的人工智能风险治理实践，引导学生探讨如何确保人工智能企业的技术和业务符合道德和</p>

						<p>法律要求，并为社会 and 用户提供可信赖和负责的解决方案。</p> <p>教学方法：讲授法</p> <p>阅读文献：</p> <p>(1) Voluntary AI Commitments,  <a href="https://www.whitehouse.gov/wp-content/uploads/2023/09/Voluntary-AI-Commitments-September-2023.pdf">https://www.whitehouse.gov/wp-content/uploads/2023/09/Voluntary-AI-Commitments-September-2023.pdf</a></p> <p>(2) Stanford Institute for Human-Centered AI: The Foundation Model Transparency Index ,  <a href="https://hai.stanford.edu/news/introducing-foundation-model-transparency-index">https://hai.stanford.edu/news/introducing-foundation-model-transparency-index</a></p>
总计	3 2	3 2				
备注 (Notes)			<p>第一章 (第 1 周), 第九、十、十一章 (第 12-16 周): 周辉</p> <p>第二至五章 (第 2-6 周): 刘灿华</p> <p>第六至十章 (第 7-11 周): 徐玖玖</p>			