**关于人工智能安全法律系统构建的探索**

**---信息安全法律法规大作业（二）**

学号： 姓名：

在新一代信息技术革命浪潮的推动下，人工智能应用越来越广泛，尤其在以大数据和深度神经网络为代表的前沿领域取得了重大创新与突破。人工智能，即能够不经人类特别干预即可在复杂多变的环境下完成任务的智能系统，能够通过机器学习不断优化决策和行动。其可以像人类一样思考、认知、规划、学习、交流或采取行动。一般而言，智能系统越能够以接近人类的方式执任务，其人工智能化的程度就越高。简而言之，人工智能实际就是创造开发出能够最大程度模拟甚或超出人类大脑智力的一套原理、技术、方法及应用系统，其最终目标是让智能机器进行各种类人化操作。当前，人工智能已成为新一轮科技革命和产业变革的重要驱动力量。国家“十四五”规划明确指出，把“人工智能”规划置于需要国家重点扶持突破发展的九大科技前沿领域任务之首，要大力培育壮大人工智能产业。特别值得一提的是，作为年度热词的“元宇宙”开始走上了社会产业发展舞台，在元宇宙的世界里，人工智能也将出演重要角色，为元宇宙赋予智能的“大脑”以及创新的内容。在人工智能被格外重视并一路绿灯快速应用、快速发展的同时，它也给我们带来了诸多的风险和挑战，其安全问题也逐渐被人们重视起来，因此政府与社会各界有必要认真地思考并形成健全的人工智能安全法律制度体系，推动人工智能安全的法制化进程。

1. **国内外人工智能安全法现状**

人工智能是一种通用技术，可以应用到从农业、制造业到服务业的各类产业，以及人类活动的各种场景之中，因此人工智能的安全十分重要。我们所理解的人工智能安全的内涵主要包括两方面。一方面是人工智能在安全领域的应用，如人工智能在越来越多的危险行业辅助或替代人类，这属于应用人工智能来保护人类安全。另一方面是人工智能自身的安全问题，这是我们重点关注的问题。如何促进人工智能的应用更加安全和道德一直是人类长期思考和不断深化的命题。现阶段的人工智能技术不成熟性会导致诸多安全风险，这也是人工智能自身安全问题出现的根源。因此制定一部总括性的、整体层面的“人工智能安全法”有其现实的需求。而国内外也都已经对人工智能安全方面的立法采取了行动：

* **国际现状**

以美国和欧盟为例。

美国极其重视人工智能安全相关问题。据统计，2016-2020年，五年之内美国至少发布了10条与人工智能安全相关的政策、法规，其中联邦政府文件4条，国会通过法案或草案6条。美国认为人工智能对维持国家经济和国家安全十分重要。2019年2月，美国总统签署行政令，启动“美国人工智能倡议”。该倡议提出应在人工智能研发、数据资源共享、标准规范制定、人力资源培养和国际合作五个方面重点发力，强调通过人工智能来确保美国的领先地位以应对来自“战略竞争者和外国对手”的挑战。美国国防部强调人工智能对国家安全的影响，将人工智能应用于军事领域。在人工智能和国家安全方面，美国每年进行年度总结，并发布以往相关战略的更新，2019年2月更新了《2018年国防部人工智能战略摘要——利用人工智能促进安全与繁荣》，11月更新了《人工智能和国家安全》。美国国防部已针对一系列军事功能开发AI应用，平衡人工智能和自主系统研究的发展与伦理，对军事人工智能应用整合的立法或监管方面做出改变。美国积极颁布法律，监管人工智能应用，保护公民隐私和个人信息，降低伦理道德影响，规范人工智能安全问题。2017年颁布了自动驾驶相关法律，探索自动驾驶汽车安全标准，包括网络安全、隐私保护等方面的规定；2018年颁布了《2018美国人工智能安全委员会法》，成立人工智能安全委员会，统筹人工智能安全问题；2019年1月开始正式起草《技术出口管制体系框架》，限制人工智能等新技术出口外流；2019年4月提出《2019年算法问责法》法案，迫使企业研究审查自身技术中是否存在种族、性别或其他方面的偏见；2020年3月通过《2020年国家人工智能计划法案》，支持有关道德、法律、环境、安全、安保以及其他与人工智能相关的社会问题的研究和其他活动；2020年11月发布《人工智能应用监管指南草案》，提出AI十项监管原则。美国还颁布了《国家人工智能研究和发展战略计划》，并于2019年6月更新，依旧把人工智能系统安全列为战略目标之一。

欧盟一直重视人工智能伦理和隐私，在为建立可信AI而努力。从2018年发布的《欧盟人工智能》战略开始，欧盟一直重视人工智能伦理和隐私问题，确保人工智能可信《欧盟人工智能》战略提出确保欧盟具有与人工智能发展和应用相适应的伦理和法律框架。2019年初，欧盟委员会发布《关于欧洲人工智能开发与使用的协同计划》，将确保信任列入四个关键发力领域，要使欧洲成为全球人工智能开发部署、伦理道德等领域的领导者；2019年4月，欧盟委员会发布了《人工智能伦理准则可信AI伦理指南》，提出可信AI七大原则：人的能动性和监督能力、安全性、隐私数据管理、透明度、包容性、社会福祉、问责机制，确保人工智能应用合乎道德。为进一步确保人工智能安全可信，2020年2月，欧盟委员会又发布了《人工智能白皮书——欧洲卓越与可信的方法》，确定了两个主要风险领域：基本权利，包括隐私保护、个人信息保护；非歧视以及责任制度的安全性。欧盟还颁布了其他的政策和法律，设立欧盟委员会直属的欧盟人工智能高级别专家组，减少人工智能对伦理道德、个人隐私等方面的冲击，建立可信的人工智能

* **国内现状**

我国目前多以行政指导的形式对人工智能发展进行规范。如国务院办公厅在2017年以《新一代人工智能发展规划》这一文件对我国现阶段人工智能发展战略进行“三步走”的总体规划，提到“到2025年初步建立人工智能法律法规、伦理规范和政策体系，形成人工智能安全评估和管控能力。”，加快人工智能的安全应用。自此，各部委与地方单位均根据整体构想制定了针对不同领域的人工智能发展战略规划，如工信部《促进新一代人工智能产业发展三年行动计划(2018-2020年)》等，为人工智能的安全发展提供政策保障。2019年3月4日，十三届全国人大二次会议发言人表示“全国人大常委会已将一些与人工智能安全密切相关的立法项目，如《数据安全法》、《个人信息保护法》和修改《科学技术进步法》等，列入本届五年的立法规划。同时把人工智能安全方面立法列入抓紧研究项目，围绕相关法律问题进行深入的调查论证，努力使人工智能创新发展，努力为人工智能的创新发展和安全发展提供有力的法治保障。”

1. **人工智能法律关键应用分析**

当前，人工智能正处于高速发展阶段，其发展方向、发展边界尚不清晰，导致规范什么和怎么规范并无统一认识。尽管如此，人们还是自发的对人工智能的很多方面进行监管，并积极制定相关法律法规。这些法律的目的并不一定是制定规范人工智能发展的普遍原则，更多的是对涉及人工智能技术应用的具体领域的规范。根据目前的情况，人工智能安全法律法规的应用重点在以下领域：

* **人工智能产品的安全规范和致人损害的法律责任问题**

与传统执行重复任务的机器不同，人工智能拥有相对的自主性，导致智能机器人、自动驾驶汽车等人工智能产品的安全性和致人损害的法律责任认定等问题极为突出。针对人工智能产品的安全性，需要从根源上探究、识别与规避人工智能产品致人损害的内在缘由，并加以规范。如为确保人工智能系统的安全，需要考虑针对其特性制定专门的检测制度；考虑到人工智能的自主能力，需要制定专门的安全规则和程序，这些规则应侧重于披露潜在风险，确保检测程序到位以便进行自我监控，或者通过实施最低安全标准来发挥规范作用。此外，在人工智能算法高深复杂、侵权过程难以查明的状况下，对于生产设计者施以道德规制将有助于避免或降低算法产生缺陷的可能。如为解决自动驾驶汽车领域的经典问题——当事故发生时，自动驾驶系统的选择程序以乘客为重还是以过马路的行人为重，一群机器学习专家设计了一个“基于投票的系统”，邀请130万人对各种处理方案进行评估，经汇总后，将这些人的整体观点应用于一系列可能的车辆行驶规则中。这使得设计者能够在自动驾驶系统的算法中引入道德决策，将公众偏好考虑在内。

关于人工智能产品致人损害的法律责任问题，涉及一系列复杂的责任认定。比如道路交通事故的认定，其归责事由只有结果的“对与错”，而无主观上的“故意”或“过失”，这就需要对过错行为的原因进行具体分析:究竟是产品的技术漏洞，还是驾驶人员的不当使用？在无人驾驶汽车造成的事故中，还包括算法引起的控制问题，甚至可能是人工智能系统超越原控制方案的“自身行为”，这会导致算法的运营商可能会负有产品责任。此外，在涉及多个责任主体时，责任的分担问题也十分突出。比如，美国Uber公司在亚利桑那州测试其自动驾驶汽车时发生了事故，测试得到州政府的支持，测试除涉及优步公司外，还包括软件开发商、汽车制造商、车上的测试员等，各自承担什么责任面临许多需要解决的法律问题。

* **数据安全与隐私保护问题**

人工智能的大多数应用需要“大数据”支持。在云计算和快速增长的数据量推动下，人工智能已在越来越多的领域得到应用。为执行任务，人工智能要大量地搜集、存储和分析数据。从数据隐私的角度来看，人工智能对数据的大量使用，使得妥善保护数据变得越来越重要。随着人工智能系统越来越多地被整合到基础设施、生产制造、商业运营和日常生活中，在医疗、交通、金融、娱乐、购物、执法等各个领域得到应用，特别是与物联网和相关的生物物联网的相互融合，关于个人的基因、面孔、财务、偏好等数据将无处藏匿。无所不在的数据捕获和优化对数据隐私和安全构成了威胁。

人工智能的核心是优化数据，作为其基础的机器学习通过“贪吃”数据来改进自身算法，数据量越大，效果越好。隐私权的保护给人工智能的发展提出了法律难题，现有的法律尚难适应人工智能发展的要求，个别法学专家甚至认为“大数据”从根本上就与数据保护相左。因此，需要在发展人工智能与数据保护之间寻求平衡。人工智能的深入发展，亦将是隐私安全风险增大与法律保护强化的过程。特别是在当前，用户数据日益集聚和掌握在大型互联网或从事数据处理的企业手中，对其保护个人隐私的责任约束就显得非常重要，因此，一个关于人工智能应用中数据保护的法律就显得格外必要。

* **算法安全性问题**

数据、算法和算力三要素助推了人工智能技术和产业的快速发展，在诸如语音识别、图像识别和自然语言处理等领域中都取得了一些突破性的进展。然而人工智能在为经济社会发展注入活力的同时，也不可避免地带来新的技术和社会安全风险，尤其是数据和算法面临着先天的安全挑战。

在算法安全方面，新一代人工智能的核心技术是通过机器学习来模仿人类智能，以实现复杂的决策过程，这无疑会存在先天性的算法黑箱与不可解释性问题。对于深度学习网络目标函数定义的不准确、不合理或者不正确也可能会导致错误甚至伤害性的结果。错误的目标函数、计算成本太高的目标函数、表达能力有限的网络都可能使网络产生错误的结果。例如，2018年3月，Uber自动驾驶汽车因机器视觉系统未及时识别路上突然出现的行人，导致与行人相撞，导致行人死亡的事故。

算法存在偏见与人工智能的不可解释性也是一个重大的问题。本质上讲，人工智能作为一种技术没有所谓的偏好，但其依赖于数据和算法等，受到使用场景和先入为主信息的影响，会导致偏见。机器就像孩子一样，不会质疑所给出的数据和算法，而只是寻找其中的模式。如果数据、算法在一开始就被曲解，那么其输出的结果也将反映出这一点。之前美国利用人工智能算法预测犯罪的人，名单暴露，其中许多无辜的人被冤枉，并且被冤枉的人中大多数为黑人，并且对于决策算法中系统为什么做出这样的决策，即使它的开发者也无法给出合理的解释。从立法上讲，采取什么措施加以防范十分重要。如为人工智能系统如何决策制定最低标准，在发布和使用前采取必要的程序加以测试，采取必要的手段对其产生的结果进行监督、评估，并视情调整。

* **劳动力市场安全问题**

人们普遍认为，人工智能可能对劳动力市场产生巨大影响。随着人工智能的发展，许多工作将被取代，甚至连医生、律师、程序员、文创人员等高技能人群也将受到波及，引发失业、加深经济和社会的不平等等一系列问题。人工智能技术的大规模应用将给现有劳动者的权利保护带来冲击。这就要求立法者从劳动保护、教育培训和再就业等方面考虑，有效应对未来劳动力市场的结构性变化，防止人工智能给就业市场带来冲击。此外，随着人工智能“工人群体”的兴起，一些法律专家也从伦理规范的角度考虑，主张人工智能具有“工人”身份，并赋予其劳动权等“特定权利与义务”，这将对未来的劳动法产生影响。总之，人工智能也会有极大概率冲击劳动力市场，造成安全问题。

* **人工智能用于攻击性活动问题**

人工智能在驱动生产力变革、推动社会发展的同时，其助长攻击性行为的威力也不容小觑。

一是人工智能技术降低了信息造假的门槛。攻击者可利用基于生成对抗网络（GAN）实现对图像、音视频的生成或修改，形成逼真的“虚假信息”。恶意伪造的图片和音视频信息不仅是对公民肖像权等个人权益的侵犯，若将其用于敲诈勒索、伪造罪证等恶意活动，会严重危害人际信任关系、影响社会稳定。基于深度伪造的虚假新闻还可能对社会舆论生态造成恶劣影响，甚至威胁国家安全，例如美国曾格外担忧虚假新闻对大选造成灾难性影响，已计划采取立法及研发奖励等方式打击相关行为。

二是人工智能技术升级网络攻击破坏性。传统网络攻击中，攻击规模和攻击效率难以兼顾，而人工智能技术的应用能够实现大规模的自动化网络攻击。一方面人工智能系统能够实现恶意软件编写和分发的自动化，大大提升渗透效率，另一方面基于被感染设备构建智能僵尸网络，利用人工智能技术可实现网络内部的信息共享、智能分析和主动攻击。

三是人工智能技术军事化可构建新型打击力量。人工智能技术除了在民用领域广泛应用，也格外受到军方的青睐。越来越多的国家试图研发“致命性自主武器系统”，使用传感器组件和计算机算法来独立识别目标，能够在没有人工控制的情况下自动攻击和摧毁目标。致命性自主武器系统将颠覆传统的战争形态，对世界各国的安全构成严重威胁，其系统本身存在的不确定性因素也将造成难以预计的后果。

1. **人工智能安全法实行的困难与挑战**

人工智能在不断进步和发展的同时，法律也必须随其发展。而这也给人工智能安全法的实行带来一定的困难和挑战。安全法实行的困难主要有:

* **法律规制不完善**

此处以应对人工智能带来的数据安全与隐私保护风险的法律为例。

2017年,我国颁布实施了《中华人民共和国网络安全法》。此后,与数据安全相关的配套规定相继出台,例如《信息技术安全个人信息安全规范》《中华人民共和国网络安全法数据安全法》(以下简称《数据安全法》)《个人信息保护法》等,甚至在《中华人民共和国民法典》中,也围绕数据安全问题,明确了数据侵权行为的民事责任关系。由此可见,我国现阶段高度重视数据安全的法治工作,积极回应了当前大数据、人工智能等技术发展的法律实践需要,立法内容较之以往有了长足进步。

但进步之余，与之相关的法律仍存在以下问题，而这也是广大人工智能安全法普遍存在的问题：

首先是**立法衔接问题**。在数据保护方面，主要表现在《数据安全法》与《个人信息保护法》未统筹规定数据处理者的安全管理模式,《数据安全法》要求的是建立健全的全流程数据安全管理制度,但《个人信息保护法》则规定分类管理和技术保护制度,易引起人工智能数据安全管理的责任交叉、立法内容的重复等问题。而这种立法重复、衔接生硬的现象在人工智能安全法中普遍存在。

其次是**法律留白问题**。由于新法颁布时间不久,难免有不周延之处，尤其面对进步和迭代速度极快的人工智能技术，为了提高法律的实用性，安全法的留白问题更加凸显。这一问题比较突出地表现在《个人信息保护法》第13条、第16条以及第26条之中,三项规定均用到了“必需”一词,但怎么认定“必需”“认定必需的原则是什么?”这些问题都有待立法者进行解答。

其三是**立法可操作性不强的问题**。这主要体现在，由于对人工智能的整体认识比较模糊，导致在部分条款拟定的比较原则化难以界定。比如说《个人信息保护法》第6条规定的“最小范围”、《数据安全法》第21条规定的“数据分级分类保护”,其具体内容都不是特别明确,关于“最小范围”“数据分级”“数据分类”的含义较为模糊,立法的可操作性差。

其四是**立法碎片化问题**。由于人工智能当前体系并不完整，发展日新月异，因此安全法大多碎片且不成体系。法律文本中没有确立统一的数据管理标准、数据审核标准和数据保护标准,而是散布在不同法律规范的各章节之中,并没有形成统一、规范的人工智能数据标准体系。

* **管理体制存在问题**

人工智能安全法应用过程中管理体制属于重要的基础保障，也是法律规定中需要明确的主要内容，目前我国在人工智能安全管理体制方面具备统一性领导与各方主体相互协调的特征，能够持续性解决复杂性问题和不确定性的问题，但是在治理期间还存在一些不足之处：

其一，管理体制具有多头管理的特点，难以准确彻底性落实人工智能安全法中规定的工作。当前我国的人工智能安全治理主要是政府部门承担重要职责，而在职责分工期间很难和人工智能多种场景相互适应，无法充分彰显出行政主管部门的主体作用，缺乏一定的针对性。

其二，缺少关于技术的支撑和保障。目前国际领域中人工智能安全治理，已经开始重视复杂性特点和不确定性特点的分析，研究此类特点所引发的安全问题，而目前我国的人工智能治理管理体制中还没有创建较为完善的人工智能安全治理的技术机制和模式，缺乏现代化的治理技术，难以严格落实人工智能安全法，进而有效维护安全性。

其三，人工智能安全治理的法律法规缺乏一定的场景适应性。仍然以数据安全为例子。虽然我国在数据安全法律、个人信息保护法律、关键性信息基础设施安全保护的法律中已经明确定义了个人数据、基础设施关键数据，但是还没有细致性提出各类数据应用场景的安全规定，尤其是医疗领域、自动驾驶领域的人工智能数据具有一定的特殊性，不能和其他人工智能数据采用相同的安全治理方案，如果不能按照数据应用的各个特殊领域特点与情况编制完善的安全治理计划和规定，将会导致安全治理效果无法提升。

* **人工智能的技术进步对现有法律不断造成冲击**

伴随着人工智能技术的不断进步，其在技术领域的创新，不论是微小的还是颠覆性的，都必然会对现有的人工智能安全法律法规造成冲击，导致法律的适用性和实用性下降。

根据以往的经验来看，在人工智能发展过程中，随着其智能性的不断提高，其产生的法律问题也越来越多。比如在人工智能的智能性还处于较低水平时，我们不会考虑人工智能作为法律意义上的“自然人”的情况；但随着人工智能的智能性的提高，类似“索菲亚”这种获得“公民”身份的机器人也诞生于世。这种并无“判例”的情况必然会对法律造成冲击。

1. **建议**

* **健全完善相关法律体系**

完备的法律法规是保证人工智能安全治理的前提,推进人工智能安全立法能够促进运用公权力对人工智能数据安全风险的遏制。人工智能数据安全治理应从“立法衔接”“法律规定的周延性”“法律规范的可操作性”“立法体系化”四个方面来加强人工智能安全领域的法治建设,为安全治理工作提供法律依据和行动指南。

* **紧跟技术发展，加快相关法律的更新迭代**

立法一方面要对人工智能发展做引领,另一方面也要规制如发展目标、路径和阶段。但由于人工智能发展飞快,立法往往跟不上发展速度。前沿的技术变革和创新的商业模式带来的不确定性,也决定了相关的立法工作会有滞后性。但新生事物并非排斥法律法规的制约,相反,法律对于新兴商业模式和技术创新的有效规范和制约能更好地引导企业、行业健康有序发展。对于人工智能安全法规，要在把握技术发展大方向的前提下，定好规制方向，不断细化要求，推动人工智能的安全发展。

由于当前人工智能技术的发展已超出法律实践，需要在理论上加强研究，及时跟进。人工智能的立法既涉及对人工智能本身的认识，包括社会伦理问题，也涉及具体应用问题，还涉及数据、算法等技术问题，不同的问题归属不同的立法范畴。这就需要从总体上加以规划，在人工智能总体立法尚不成熟时，具体应用领域立法不应停滞，应成熟一项、出台一项，先易后难，逐步形成体系。

* **优化管理体制**

针对人工智能的高速发展和所产生的一系列安全问题，政府应当尽早形成以全流程监管为主导的、全方位的人工智能监管体制，以实现人工智能在安全法的架构下正常运作，并建立合理的监管机制、明晰的法律责任体系，以达到对人工智能的全面科学监督，从而更好地促进人工智能的健康发展。

比如，深化顶层设计，实现精准监管，提高管理效率，彰显出行政主管部门的主体作用，缺乏一定的针对性；制定监管标准，实施规范管理，创建较为完善的人工智能安全治理的技术机制和模式，采用现代化的治理技术，严格落实人工智能安全法；重视典型案例，建立有效举措，重视法律应用的具体情境；加强跨部门、跨行业协作，人工智能属于通用性技术，涉及面广，应用领域宽，带来的问题新，所涉及的诸多立法问题在现有的法律框架和制度体系下难以找到答案，同时，人工智能带来的很多法律问题属于跨领域问题，涉及技术标准、国家安全、伦理道理等，需要召集行业专家、法律专家和社会学家共同研究论证。因此，应加强跨部门、跨行业协作，在聚集经验智慧、形成普遍共识的基础上，推动人工智能领域的立法。

* **提高司法机关的技术手段**

人工智能环境下发生的个人信息侵权，由于手段相对隐蔽,侵权的方法复杂多样,数量众多,司法机关在应对侵权行为时面临巨大挑战。司法机关必须进一步增强现代司法技术手段,运用现代智慧云和大数据科技保障公民个人信息安全。人工智能网络侵犯程序面临的技术问题之一是网络侵权司法过程中复杂的侦查取证等环节，司法领域的相关专门的技术人员都需要通过提高自身对人工智能计算机算法网络等技术程序知识的全面熟悉,构建出专业高效的智能司法网络程序,通过技术手段提高其取证的效率。利用人工智能可以复制犯罪人的侵权过程，在合理范围内系统保护个人信息，创建个人信息保护壁垒，对收集的信息进行过滤，检查收集和处理的信息是否合法，并及时反馈。

* **加强国际协作，引领人工智能安全规范的国际治理**

人工智能发展关系未来人类共同命运，加快国际合作，共同制定发展规范是大势所趋。除应重视技术发展外，还应加强与各国的合作，积极参与相关国际规范制定。当前，国际社会普遍关注网络安全和数据隐私，以及人工智能对社会伦理的影响，要主动加强在产业、智库、学者层面的沟通交流，主动引领命题，提供中国经验和中国解决方案。此外，中国的大市场为发展人工智能提供了很好平台，在保障国家安全的前提下，积极创造条件，推动国际合作，边应用、边治理，推进人工智能安全领域的国际立法。

1. **总结**

人工智能技术发展的势头无法阻挡，人工智能技术发展的高度无法预测，但无论如何，法律规范是规范人类行为的，这一点是不容质疑的，如何设计涉及人工智能安全的相关法律规则，要以如何更好地助益于人类为出发点。人工智能的发展并没有发展到突破现有法律主体、权利、责任相关立法框架的地步，现有的安全方面的立法只要针对人工智能在安全方面带来的挑战积极做出调整就可以。但根据人工智能技术的发展趋势，其智能性将不断得到体现。我们一定要把握人工智能技术的发展势头，重视其安全发展，在人工智能安全法的框架下规范其发展，使人工智能成为人类更加强大且适用的工具。