

<b>Forum:</b>	Chinese Language Committee (High School, Year 9-12)
<b>Agenda:</b>	On measures to regulate the increasing impact of artificial intelligence 关于管制人工智能（AI）日益增长的影响的措施
<b>Student Officer:</b>	Yvonne Wang

---

## Introduction

人工智能（Artificial Intelligence），是计算机科学的一个分支，它是一门通过研究、开发之后，可用于模拟、延伸或扩展人的智能的理论、方法和技术的一门新的技术科学。人工智能的定义可以分为两部分，即“人工”和“智能”。人工智能是一个通过研究使计算机来模拟人的某些思维过程和智能行为（如学习、推理、思考、规划等）的学科，主要包括使用计算机实现智能的原理、制造类似于人脑智能的计算机，使计算机能实现更高层次的应用。同时，人工智能也涉及到了计算机科学、心理学、哲学和语言学等众多学科。在联合国教科文组织第41届大会上通过的《人工智能伦理建议书》里人工智能的定义是：“将人工智能系统视为有能力以类似于智能行为的方式处理数据和信息的系统，通常包括推理、学习、感知、预测、规划或控制等方面”。

1956年，四位年轻学者麦卡锡（McCarthy J）、明斯基（Minsky M）、罗彻斯特（Rochester N）和香农（Shannon C）共同发起和组织召开了达特茅斯（Dartmouth）夏季专题讨论会，研究用机器模拟人类智能。在讨论会上，麦卡锡提议用人工智能（artificial intelligence）作为这一交叉学科的名称，定义为制造智能机器的科学与工程，标志着人工智能学科的诞生。自1956年人工智能诞生以来，从历经艰辛与坎坷，再到取得了举世瞩目的成就，特别是在机器学习、数据挖掘、计算机视觉、专家系统、自然语言处理、模式识别、机器人等等和人工智能相关的应用都产生了良好的经济效益和社会效益。数字化、网络化和智能化是信息社会发展的必然趋势，人工智能的长期目标是建立达到人类智力水平的人工智能。人工智能是研究可以理性地进行思考和执行动作的计算模型的学科，它是人类智能在计算机上的模拟。人工智能作为一门学科，经历了孕育、形成和发展等几个阶段，并且还在不断地发展。知识表示、推理、学习、智能搜索和数据与知识的不确定性处理是人工智能的基本研究领域，人工智能的研究途径主要有以符号处理为核心的方法、以网络连接为主的连接机制方法，以及以感知和动作为主的行为主义方法等等。

随着人工智能（AI）的不断的日益增长，人工智能技术可以对人类提供较大的助益并惠及全球所有国家，但与此同时也会引发关于人工智能的根本性的伦理上的问题，例如：第一，人工智能技术可能内嵌并加剧偏见，由此可能会导致歧视、不平等、数字鸿沟和排斥等一系列问题，并对文化、社会和生物多样性构成威胁，造成社会或经济鸿沟；第二，人工智能技术对于社会的多方面会存在潜在影响，包括但不限于

于人的尊严、人权和基本自由、性别平等、民主、社会、经济、政治和文化进程、科学和工程实践、动物福利以及环境和生态系统。

## Key Terms

**大数据 (big data)** - 大数据是一个不断变化的目标，具有Volume(量大)、Velocity(速度快)、Variety(多样性)、Value(价值)四个显著特点。同时，大数据技术也为人工智能提供了强大的存储能力和计算能力。

**云计算** - 云计算从技术体系结构上来看，云计算可以看成是人工智能的一个基础，可以为人工智能提供算力支撑。

**数据库** - 数据库系统和人工系统的互相融合，具有一定的互补性。这两者的互相融合，可以有效实现数据库的智能检索和查询优化功能，并且也可以利用当前的人工智能将数据库当中的大规模数据进行整合归纳，促使人们在使用数据库技术的时候可以更加便携有效，提高数据库的智能性。

**数字鸿沟** - 数字鸿沟，是指在全球数字化进程中，不同国家、地区、行业、企业、社区之间，由于对信息、网络技术的拥有程度、应用程度以及创新能力的差别而造成的信息落差及贫富进一步两极分化的趋势。

## General Overview

随着人工智能现如今的高速发展，人工智能技术将会代替信息技术成为社会的主导技术，推动社会进入一个新的崭新的形态—智能化社会。与此同时，随着人工智能算法的创新将从现在的专用智能向通用智能转化，智能系统的智力将得到很大提升，解决问题范围更广，质量更高。通过智能技术发展，人们的生活将更加轻松，比如说大多数服务行业以及一些繁重的工作（客服，流水线工人，快递员等等），都将会被无人系统所代替，通过算法的开发，无人系统将会代替人们完成很多的工作。

但随着人工智能应用的广泛的普及，人类对人工智能的依赖性越来越强。这导致了人类开始逐渐缺乏自主性、创造性与责任感。虽然人工智能可以通过大数据挖掘，算法预测和算力基础上的分析，给人类生活带来便捷。但同时也会对人类社会信息的安全性产生挑战。

### 澳大利亚

澳大利亚对于人工智能对企业成功的战略重要性普遍持积极肯定的态度。79%的受访者认为人工智能在两年内将对其企业“非常”或“至关”重要。尽管如此，有50%的受访者表示人工智能正在帮助他们“追赶”或“紧跟”竞争对手的步伐，而非确立显著的领先优势。

## 加拿大

加拿大在人工智能上采取较为谨慎的策略，仅25%的加拿大人工智能应用者表示他们目前已在其产品和服务中运用人工智能，这属于较低水平。在国家层面上，加拿大正致力于推动人工智能能力的协同发展。尤其在人才方面，政府制定了多项政策促进移民，为具备人工智能相关技术能力的人才提供更为便捷和开放的移民机会。

## 欧盟

在2016年4月14日，欧洲议会投票通过的《一般数据保护条例》（General Data Protection Regulation）里要求AI公司手动审查重要的算法决策，在用户使用了某项服务或软件后，如果认为该服务或软件收集了个人数据并侵犯隐私，就可以要求公司立刻删除，保护个人隐私。

2021年4月21日，欧盟委员会公布了名为“Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts”的草案，对人工智能技术应用进行了风险评定。根据这份草案，欧盟将全面禁止大规模监控和利用人工智能技术的社会信用体系，同时对特定领域的“高风险”应用进行严格限制。该草案目的在于为实现可信赖的人工智能生态系统提供牢固的法律框架。该草案将人工智能应用分为了完全不可接受风险、高风险和低风险等不同等级，其中有三种应用被完全禁止。第一种是在不知不觉中对人类意识进行操控，从而影响其决定或扭曲其行为，进而对人类造成身体或心理的伤害。第二种情况是利用儿童或残疾人的脆弱性对其造成伤害。第三种则是基于人工智能系统上的社会信用体系。

## 日本

日本在2013年修正的《个人信息保护法》中，制定了规则来强化个人信息保护，要求企业履行尽可能消除客户个人数据的义务，并规定未经当事人同意，不得向其他人提供数据等法律来保护个人数据的安全。

## Timeline of Events

### *Date*

人工智能的孕育期（1956 年以前）

### *Description*

人工智能的孕育期大致可以认为是在 1956 年以前的时期。这一时期的主要成就是数理逻辑、自动机理论、控制论、信息论、神经计算和电子计算机等学科的建立和发展，为人工智能的诞生，准备了理论和物质的基础。

## 人工智能的形成期（1956～1969 年）

人工智能的形成期大约从 1956 年开始到 1969 年。这一时期的主要成就包括 1956 年在美国的达特茅斯（Dartmouth）大学召开的为期两个月的学术研讨会，提出了「人工智能」这一术语，标志着这门学科의 正式诞生；还有包括在定理机器证明、问题求解、LISP 语言、模式识别等关键领域的重大突破。

## 低潮时期（1966～1973 年）

人工智能快速发展了一段时期后，遇到了很多的困难，遭受了很多的挫折。如鲁宾逊的归结法的归结能力是有限的，证明两个连续函数之和还是连续函数时，推了十万步还没有推出来。人们曾以为只要用一部字典和某些语法知识即可很快地解决自然语言之间的互译问题，结果发现并不那么简单，甚至闹出笑话。如英语句子：“The spirit is willing but the flesh is weak”（心有余而力不足），译成俄语再译成英语竟成了：“The wine is good but the meat is spoiled”（酒是好的，肉变质了）。这里遇到的问题是单词的多义性问题。那么人类翻译家为什么可以翻译好这些句子，而机器为什么不能呢？主要原因在于翻译家在翻译之前首先要理解这个句子，但机器不能，它只是靠快速检索、排列词序等一套办法进行翻译，并不能「理解」这个句子，所以错误在所难免。

## 基于知识的系统（1969～1988 年）

1965 年，斯坦福大学的费根鲍姆和化学家勒德贝格（Lederberg J）合作研制出 DEN-DRAL 系统。1972～1976 年，费根鲍姆又成功开发出医疗专家系统 MYCIN。此后，许多著名的专家系统相继研发成功，其中较具代表性的有探矿专家系统 PROSPECTOR、青光眼诊断治疗专家系统 CASNET、钻井数据分析专家系统 ELAS 等。为了适应人工智能和知识工程发展的需要，日本在 1981 年宣布了第五代电子计算机的研制计划。其研制的计算机的主要特征是具有智能接口、知识库管理和自动解决问题的能力，并在其他方面具有人的智能行为。

### 神经网络复兴 (1986 年至今)

20 世纪 80 年代神经网络复兴的真正推动力是反向传播算法的重新研究。该算法最早由 Bryson 和 Ho 于 1969 年提出。1986 年，鲁梅尔哈特 (Rumelhart D E) 和麦克莱伦德 (McClelland J L) 等提出并行分布处理 (Parallel Distributed Processing, PDP) 的理论 [Rumelhart et al. 1986]，致力于认知的微观结构的探索，其中多层网络的误差传播学习法，即反向传播算法广为流传，引起人们极大的兴趣。世界上许多国家掀起了神经网络研究的热潮。从 1985 年开始，专门讨论神经网络的学术会议规模逐步扩大。1987 年在美国召开了第一届神经网络国际会议，并发起成立国际神经网络学会 (INNS)。

### 智能体的兴起 (1993 年至今)

20 世纪 90 年代，随着计算机网络、计算机通信等技术的发展，关于智能体 (Agent) 的研究成为人工智能的热点。1993 年，肖哈姆 (Shoham Y) 提出面向智能体的程序设计 [Shoham 1993]。1995 年，罗素 (Russell S) 和诺维格 (Norvig P) 出版了《人工智能》一书，提出「将人工智能定义为对从环境中接收感知信息并执行行动的智能体的研究」 [Russell et al. 1995]。所以，智能体应该是人工智能的核心问题。斯坦福大学计算机科学系的海斯-罗斯 (Hayes-Roth B) 在 IJCAI' 95 的特约报告中谈到：「智能体既是人工智能最初的目标，也是人工智能最终的目标」 [Hayes-Roth 1995]。

## UN Involvement, Relevant Resolutions, Treaties and Events

联合国教科文组织(The United Nations Education, Scientific and Cultural Organization)与世界科学知识与技术伦理委员会(World Commission on the Ethics of Scientific Knowledge and Technology)在联合发布的报告(2016年)主要讨论了机器人的制造和使用促进了人工智能的进步，以及这些进步所带来的社会与伦理道德问题。

具体分为以下四个方面：

- 自动化机器人的使用带来的挑战

在2016年，欧洲议会发布的关于人工智能和机器人的报告，表达了其对于机器人将给人类带来的风险的关注，包括安全、隐私、诚信、尊严、自主。为了应对这些风险，欧洲议会讨论了未来可能面对的道德挑战以及应对的监管措施。

- 机器人技术与机械伦理学

联合国教科文组织与世界科学知识与技术伦理委员会的报告(2015年)认为“机械伦理学”的大部分领域仍然处于没有规范的状态，一方面是因为政府无法跟飞速的科技发展相同步，另一方面是因为“机械伦理学”的复杂性和其无法预知的本质。

- 迈向新的责任分担机制

联合国教科文组织与世界科学知识与技术伦理委员会的报告(2015年)探讨了一个复杂的问题，即在制造一个机器人需要不同的专家和部门合作的情况下，谁该为机器人的某次操作失灵负有责任。报告认为，关于机器人伦理的考虑，不应该局限于某次事故或者失灵造成的人身损害，更应该包括智能机器人带来的心理伤害，如机器人侵犯人的隐私或人因为机器人的类人行为而对其过分依赖等。报告提出了一个可行的解决办法，即采取责任分担的解决途径，让所有参与到机器人的发明、授权和分配过程中的人来分担责任；另一个解决办法就是让智能机器人承担责任，因为智能机器人确实拥有前所未有的自主性，并且拥有能够独立做出决策的能力。

- 决策可追溯的重要性

联合国的报告认为，在对机器人及机器人技术的伦理与法律监管中，一个至关重要的要素是可追溯性，可追溯性的确立才能让机器人的行为及决策全程处于监管之下。

接着，在2021年联合国教科文组织第41届大会上获得通过的《人工智能伦理建议书》是全球首个针对人工智能伦理制定的规范框架。该建议书旨在让人工智能系统可以造福人类、个人、社会、环境和生态系统，同时防止危害。它还旨在促进和平利用人工智能系统。目的是在全球现有人工智能伦理框架之外，再提供一部全球公认的规范性文书，不仅注重阐明价值观和原则，而且着力于通过具体的政策建议切实落实这些价值观和原则，同时着重强调包容、性别平等以及环境和生态系统保护等问题。

同样，经合组织（OECD）也高度重视AI准则的制定，于2019年5月发布首套政府间人工智能政策指南《OECD人工智能原则》，共42个国家签署，随后也被G20采纳。

## Possible Solutions

阐明人工智能的明确法律基础和依据

应制定合适的管控和发展人工智能的法律，并在此基础上，让在所有公共部门机构在将人工智能应用于公共服务之前，都应发布一份声明，来阐述及说明他们对人工智能的使用如何符合相关的法律法规。

#### 设立监管保证机构

鉴于目前人工智能的发展和实施速度过快，建议设立一个监管保证机构，以确定监管环境中的人工智能使用和发展差距，并就与人工智能相关的问题向各个监管机构和政府提供建议。

#### 影响评估

应考虑如何将人工智能影响评估要求纳入现有流程，以评估人工智能对公共标准的潜在影响。此类评估应是强制性的，并应予以公布。

#### 评估公共标准的风险

应用及开发人工智能的公共服务提供商（包括公共和私人）应在项目设计阶段时评估并拟议人工智能系统对公共具有的潜在影响，并确保该系统的设计减轻了任何已确定的标准风险。并在以后每次对人工智能系统设计进行重大更改时，都需要进行标准审查。

#### 确认责任，监测和评估

第一、公共和私人人工智能服务提供商应明确分配和记录人工智能系统的责任，并确保人工智能系统运营商能够以有意义和不违法的方式履行其责任。

第二、公共服务提供商（包括公共和私人）应监控和评估其人工智能系统，以确保其始终按预期运行。

第三、公共服务提供商（包括公共和私人），应建立监督机制，允许对其人工智能系统进行适当审查。

## Bibliography

“Future in the Balance? How Countries Are Pursuing an AI Advantage.” *Deloitte Insights*, [www2.deloitte.com/insights/stateofAI-global](http://www2.deloitte.com/insights/stateofAI-global). Accessed 3 Dec. 2022.

“Artificial Intelligence | United Nations - CEB.” *Unsceb.org*, [unsceb.org/topics/artificial-intelligence](http://unsceb.org/topics/artificial-intelligence).

“论文在线阅读—中国知网.” *Kreader.cnki.net*, [kreader.cnki.net/Kreader/CatalogViewPage.aspx?dbCode=CAPJ&filename=JXSH20221028000&tablename=CAPJLAST&compose=&first=1&uid=WEEvREcwSIJHSldSdmVpa3VEenlwUxDTGpseFJSMmNiQ1NMRDhuN0dhdz0=](http://kreader.cnki.net/Kreader/CatalogViewPage.aspx?dbCode=CAPJ&filename=JXSH20221028000&tablename=CAPJLAST&compose=&first=1&uid=WEEvREcwSIJHSldSdmVpa3VEenlwUxDTGpseFJSMmNiQ1NMRDhuN0dhdz0=). Accessed 3 Dec. 2022.

“Challenges of AI.” *Chatham House – International Affairs Think Tank*, 22 Mar. 2022, [www.chathamhouse.org/2022/03/challenges-ai](http://www.chathamhouse.org/2022/03/challenges-ai).

*Zhihu.com*, 2018, [www.zhihu.com/pub/book/119642381](http://www.zhihu.com/pub/book/119642381). Accessed 3 Dec. 2022.

*Baidu.com*, 2019, baike.baidu.com/item/.

*Unesco.org*, 2022, unesdoc.unesco.org/ark:/48223/pf0000381137.

*Baijiahao.baidu.com*, baijiahao.baidu.com/s?id=1718744465798045163&wfr=spider&for=pc. Accessed 7 Dec. 2022.

*Baijiahao.baidu.com*, baijiahao.baidu.com/s?id=1717586861659214975&wfr=spider&for=pc. Accessed 7 Dec. 2022.

*Kns.cnki.net*, kns.cnki.net/kcms/detail/detail.aspx?dbcode=CAPJ&dbname=CAPJLAST&filename=XTBZ20221123002&uniplatform=NZKPT&v=iZHdjExSp0cueTAAhSvp3UjQte06\_L\_aC8A7ToLK6hUnlrXY2NSNnXzRGYD07J8W. Accessed 7 Dec. 2022.

*Kns.cnki.net*, kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDAUTO&filename=GDKJ202210018&uniplatform=NZKPT&v=2NAVdnXAcp3yodZMSbcJhyv96CwNZtn8\_IGyCEPG9a2yCVD-1lxDMPeldGkG4QJy. Accessed 7 Dec. 2022.

*Kns.cnki.net*, kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFDLAST2022&filename=GKFC202209012&uniplatform=NZKPT&v=471xm3uukeMLV6liIaPNJOEHKx0kDOqND-4Z\_6nXwGDnqQqu2ydN4Pjl0UbH6WW. Accessed 7 Dec. 2022.

*Fx361.com*, 2021, www.fx361.com/page/2021/0916/8859502.shtml. Accessed 7 Dec. 2022.

*M.haiwainet.cn*, m.haiwainet.cn/middle/348217/2020/1231/content\_31951912\_1.html. Accessed 7 Dec. 2022.

## Contact Information:

President Chair: Yvonne

WeChat ID: ylzc\_yone

Email: 23ywang@student.uiszc.org

Rosie Lee – Secretary General

[23rlee@student.uiszc.org](mailto:23rlee@student.uiszc.org)

Peter Pang – Head of Chair

[24ppan@student.uiszc.org](mailto:24ppan@student.uiszc.org)

Luke Ross Nuttall – Director of U2NESCO

[lrn@uiszc.org](mailto:lrn@uiszc.org)