



The Holy See

教宗方济各

第 57 届世界和平日文告

2024 年 1 月 1 日

人工智能与和平

新年，是上主赐给我们每一个人的恩宠时刻，在这新的一年开始，我愿向全体天主子民，向不同的民族、国家及政府首长、不同宗教和民间社团的领袖，以及我们这个时代的所有男士、女士致意，献上我对和平的热切祝愿。

1. 科技的进步作为走向和平的途径 圣经里证实天主赐下圣神给人类，使他们有「智能、技能和知识，能做各种工程」（出卅 五 31）。人类的聪明才智是造物主赐予尊严的表现，因祂按自己的肖像和模样创造了我们（参阅：创一 26），且让我们能有意识并自由地回应祂的爱。科技则以特别的方式，彰显了人类智慧基本的相关特质——人类创造潜力的杰出成果。

梵二大公会议文献《论教会在现代世界》牧职宪章里重申此一真理，宪章提到「人类经常劳动并动用脑筋，以改善自己的生活。」¹ 当人类藉助技术，努力使大地成为适合全人类的住所时，² 他们就履行了天主的计划，同时与祂的旨意合作，使这创造臻于完美，并促成各民族之间的和平。科学技术的进步，已使人类社会更加有序，并更具有兄弟情谊般的共融与自由，导致了人类的进步和世界的变革。

由于科学与技术上令人瞩目的成就，过去使人类生活饱受煎熬且造成极大痛苦的疾病，如今得以医治，我们理当为此高兴与感激。科技的进步，使得人类能控制现实的程度前所未有，于是我们可以有大量的选择，包括那些对我们的生存和家园会造成威胁的选择。³

新的信息科技，特别是在数字领域的惊人进步，给了我们令人兴奋的机会，也带来重大的风险，对于追求人民之间的正义与和谐产生了严重的影响，因此必须提出许多迫切的问题：这些新的数字科技会带来那些中、长期的后果？对个人的生活和社会、国际的稳定及和平，会有什么影响？

2. 人工智能的未来：在承诺与风险之间 近数十年来，信息科技和数字技术的进步，已使全球社会及各种动态产生深远的改变。新的数字工具甚至正在改变传播、公共行政、教育、消费、人与人之间的互动，以及我们日常生活中无数其他层面的样貌。

此外，从遍布因特网的数字足迹中，科技利用各种算法来提取数据，使它们能在我们不知情的情况下，控制我们的心理和相关习惯，以达到它们的商业和政治目的，因而限制了我们有意识的选择自由。在承载过量的网络空间中，他们可以经常在使用者察觉不到的情况下，根据使用者的选择标准来建立数据流。

我们必须记住，科学研究和技术创新不是脱离实体的，也不是「中立」⁴的，总会受到文化的影响。作为全体人类的活动，它们所发展的方向反映出在任一特定年龄的个人、社会和文化价值观之下所作的选择。人类以什么方式对待周遭世界，一如所产生的结果，是有道德的层面，是与人所作的决定密切相关，意即设计实验计划，其结果是针对一些特定的目标。

用在人工智能的形式也是一样。至今为止，在科学和技术领域中，对于人工智能，还没有一个单一的定义。这个词汇本身如今已成为日常生活的用语，包含各种科学、理论和技术，其目的是使机器的功能可以复制或模仿人类的认知能力。我们用复数谈论「不同形式的智能」，更有助于强调这些系统与人类之间不可逾越的差距，不论那智能是多么令人惊叹和强大，最终也只是「零碎」的，意思是说，也只能模仿或重现人类智能的某些功能。同样，多种的智能形式也指出以下的事实，那就是这些设备之间存有极大的差异，应该始终被视为是一种「社会技术系统」，因为任何人工智能设备的影响——不论其基础技术如何——都不仅取决于技术设计，也取决于其拥有者及开发者的目标和利益，以及应用的场景。

为此，人工智能应该被理解为恍如星系的不同现实。我们不能事先假定它的发展会有益于人类的未来及人民之间的和平。只有在我们展现自己有能力负责，并尊重「包容、透明、安全、公平、隐私和可靠」⁵等人性基本价值的情况下，才有可能得到这正面的结果。

同样，我们也不能只是假定那些设计算法和数字科技的人，愿意承诺对他们的行为负起责任且是合乎道德的。我们必须加强，或有必要时去建立机构，来负责审查这一领域中所出现的道德问题，并保护那些采用人工智能或受其影响者的权益。⁶

因此，科技的巨大扩张，必须随之有适当的培育，使科技对未来的发展负责。只要人类一旦屈服于自私自利、追求利润和渴望权力等，自由与和平的共存就受到威胁。为了个人及团体的整体发展，我们有责任拓宽视野，把科学技术研究导向追求和平及共同利益层面。⁷

每个人与生俱来的尊严，以及把我们凝聚成一个人类大家庭的兄弟情谊，必须成为新技术发展的根基，并在这些技术被应用之前，作为毋庸置疑的评估标准，这样，数字的进步才能与正义并行，且有助于和平大业。技术发展若不能提升全人类的生活质量，反而加重不平等及冲突，则永远不能视之为真正的进步。⁸

人工智能会越来越重要。它造成的问题既是技术上的，也是人类学、教育、社会和政治上的。例如

，它保证我们免于辛苦的工作，提高制造效率、更便捷的交通运输、更方便 的市场，且大大革新了收集、整理和确认数据的过程。然而，我们必须小心留意如今的 快速转变，并要谨慎加以管控，好使基本人权得以保护，同时强调推动人类整体发展的 制度及遵守法律。人工智能应有助于发展我们最佳的潜力和最崇高的志向，而不是与之 相抗衡。

3. 未来的技术：能自行学习的机器 人工智能有多种形式，而基于「机器学习」(machine learning) 所发展的人工智能，虽 然还在开创阶段，但已经给社会的结构带来可观的改变，对文化、社会行为及建立和平 等也产生了深远的影响。

像这种机器学习或「深度学习」(deep learning) 的发展，引发了超越技术和工程领域的 问题，事关深刻了解生命的意义、知识的建构及心灵获得真理的能力。

例如，某些仪器虽有产生句法和语意一致的文本，但不能保证它的可靠性。据说它 们是在「幻想」，也就是说，会创造出一些乍看之下似乎可信，其实是毫无根据或有不 实偏见的陈述。当人工智能用于传播假新闻，引起人们对媒体的日渐不信任，就会带来 严重的问题。隐私、数据所有权及知识产权等，也是此一技术会引起风险的领域。我们 还可以加上这些技术的滥用所带来的其他负面后果，诸如：歧视、干预选举、对社会的 监控增加、滥用因特网排斥他人，以及与社会越来越脱节的个人主义等问题，日益严 重。这种种因素都会助长冲突，阻碍和平。

4. 在技术官僚范例中对限度的意识 我们的世界太广大、变化多端且复杂，永远无法充分了解和完整分类，即使借助最先进 的算法，人类的思想也永远无法将世界里丰富的东西一一探讨。这样的算法无法保证对 未来的预测，只能提供统计上的近似值。不是所有事都能预测，也不是一切都能计算；「现实终究大于思想」，⁹ 不论我们的计算能力有多惊人，总是会有一些无法计量的空 间，让我们对任何量化的尝试，力有未逮。

此外，人工智能所分析的庞大数据，本身不能保证客观公正。用算法推断信息时，总是 存在着失真的风险，复制了原来环境中的不公正与偏见。算法越快速、越复杂，就越难 了解为何会产生某一特定的结果。

「智能」机器会以越来越高的效率，来执行指派给它们的任务，但它们操作的目的和意 义，仍然会由人类自己拥有的宇宙价值观来决定或操控。其风险就是，在某些决定背后 的标准，会变得不清晰，对那些决定的责任被隐匿，制造商则能规避他们为小区利益而 行动的义务。在某种意义上，这是技术官僚制度所青睐的，将经济与技术连在一起，并 特别重视效率，而忽视与直接利益无关的任何事物。¹⁰

这应该促使我们省思在当前技术官僚和效率导向的心态上往往被忽略的一些事，这对个 人及社会发展极为重要，那就是「对限度的意识」。按定义来说，人必有一死。但人意 图借着科技来克服每一种限度，不能抑制地渴望控制一切，我们就可能失去对自己的掌 控；为了追求绝对的自由，我们可能陷入「技术独裁」的漩涡。承认并接纳我们身为受 造物，即承认我们是有限度的，这是我们得到，或更好说是欣然接受「成就」这份恩赐 所不可或缺的条件。在技术官僚的范例脉络中，被普罗米修斯般的心态激发，也就是在 自以为是的意识型态下，不平等的现象可能会急遽成长，知识和财富

集中在少数人手中，使民主社会及和平共存受到严重的危害。11

5. 伦理上的急迫问题 未来，申请抵押贷款者的信用、申请工作者的适用性、罪犯再犯的可能性，或是接受政治庇护或社会福利等，都可由人工智能系统来决定。这些系统的引进，由于缺乏不同层级的调停，特别容易造成各种形式的偏见及歧视：系统性错误很容易倍增，不但在个别案例中会产生不公正，而且由于骨牌效应，也会产生真正的社会不平等。

有时人工智能的各种形式会受制于事先确定的、与激励和劝阻相关的选项，或透过以信息设计来操作的系统来规范人的选择，这些似乎都能影响个人的决定。这些形式的掌控或社会规范，需要非常谨慎地去关注及加以监督，同时也意味着，制造商、部署者和政府当局都需要负起明确的法律责任。

对于自动化过程的依赖，例如：广泛使用监视器，或采用社会信用体系，将个别人士分类，同样会对社会结构产生深远的恶果，因为它在人民之间造成了阶级之分。这种人工分类的过程，也会导致权力冲突，因为它不仅关系到虚拟用户，也关系到真实的人民。基于对人类尊严的基本尊重之要求，我们拒绝将人的独特性视为等同于一组数据。算法必不可用来决定我们如何了解人权，从而把人类的固有价值，如：同情、仁慈和宽恕置之一旁，而且也无法避免一个人有可能用它改变并摆脱自己的过去。

在这种情况下，我们也不能不考虑到新科技对工作场所的影响。过去专属于人力劳动领域的那些工作，如今迅速被人工智能在工业上的应用所取代。在这方面，也存在着少数人获得极大利益，而多数人却付出贫困代价的风险。尊重劳工的尊严，以及就业对个人、家庭和社会经济福祉的重要性，应该是国际社会最应优先关注的，因为这些科技形式已更深入地渗透我们的工作场所中。

6. 我们要化刀剑为犁头吗？在这些日子，当我们环顾这世界，会看到一个无法逃避的严重道德问题，与军火产业有关。利用遥控系统指挥军事行动的能力，会使人疏于留意系统所造成的破坏，以及对使用武器负责的心理负担，以致对战争的巨大悲剧更加冷漠无感。对于正在兴起的所谓「致命自动武器系统」，包括武器人工智能化的研究，会引起对道德伦理的严重挑战。自动武器系统绝不可能负起伦理道德的责任，唯有能力作道德判断及合乎伦理的决定，这是一组复杂的算法远不能及的，毕竟那能力不能简化为将机器程序化，不论那机器多有智能，终究是一部机器。因此，人类亟需对武器系统进行充分、持续且有意义的监督。

我们也不能忽视尖端武器最后落在不当人士手中的可能性，例如：促成恐怖分子的攻击或颠覆政府的合法体制。总之，这世界不需要有助于商业及武器交易的「新」技术，这只会导致不公义的发展或疯狂的战争。若这样做，不只是人的智慧，甚至人类心灵本身也有可能变得更「人工化」。最先进的科技，不应用来助长以暴力解决冲突，而应该为和平铺路。

从更积极的角度来看，如果将人工智能用于促进整体人类的发展，就可以在农业、教育和文化上带来重要的新发明，改善整个国家和人民的生活水平，并增进人类的手足情谊及社会的友爱。最后，我们如何以最小弟兄姊妹、易受伤害者以及最需要帮助者为念，来使用人工智能，才是衡量人性的真正标准。

一个真正人性化的观点，以及对我们世界有更美好未来的渴望，无疑表示我们需要进行跨领域的交谈，以实现算法合乎伦理——即演算伦理——的发展，其价值观会塑造出新科技的走向。¹²在研究工作的起始，即应考虑到伦理问题，并且持续到实验、设计、生产、分发及营销的各个阶段。这就是设计伦理学的做法，也是教育机构和决策者必须扮演的重要角色。

7. 教育面临的挑战 一个尊重人类尊严并为其服务的技术发展，对我们的教育机构和文化世界有明显的影响。数字科技大幅增加了传播的机会，使我们能以新的方式彼此相遇。然而，我们仍必须不断地省思，这科技会引导我们进入何种关系。我们的年轻人成长于科技普及的文化环境，这也不可避免地对我们教学、教育和培训的方法带来了挑战。

使用各种形式的人工智能来从事教育，其首要目标应该是促进批判性思考。各个年龄层的使用者，特别是青年，对于在网上收集的数据，或由人工智能系统生成的内容，必须培养一种辨明的能力。学校、大学和科学社团都面临挑战，要帮助学生和专业人士理解科技的发展和应用上的社会层面及伦理道德。

训练人们使用新的传播工具，不但要考虑错误讯息和「假新闻」，也要考虑到下面令人困扰的情形复发，那就是「隐藏在新技术背后且不断蔓延的由来已久的恐惧……。」¹³令人遗憾的是，我们发现自己要再次面对「建立围墙文化的诱惑，竖立围墙，以阻止我们与其他文化和与其他人民交流」¹⁴的恶果，以及对和平友好共存发展的影响。

8. 发展国际公约所面临的挑战 从人工智能的全球化规模可以清楚地看到，除了主权国家有责任由内部规范它的使用外，国际组织在达成多边协议及其应用和执行方面，也有决定性的角色要扮演。¹⁵关于这方面，我力促各国的全球社群能共同合作，以制订一个有约束力的国际协议，来规范各种形式的人工智能的发展和运用。当然，规范的目的不应该只是防止有害的使用，也要鼓励做最好的用途，要能激发新颖且有创意的做法，并鼓励个人或团体的创新。¹⁶

总而言之，在寻求能对数字科技的开发者提供道德指导的规范模型过程中，必须要能指出人类的价值观，才能驾驭社团制订、采用并执行所需要的规范架构。为产生各种形式的人工智能的工作而制订的道德指南，不能不思考人类存在意义、保障基本人权和追求正义与和平等更深刻的议题。伦理和法律上的分辨，可成为一个可贵的机会，来共同省思科技在个人和公共生活中应扮演的角色，以及要如何应用才能有助于创造一个更公正、更人道的世界。因此，在谈论人工智能的规范时，所有利害关系人的声音都应纳入考虑，包括那些在全球决策过程中经常被忽视的贫困者、无权无势者及其他人士。

* * * * *

我希望上述的反思，能鼓励大家努力，确使人工智能形式的进展最终将成为服务于人类友谊与和平的大业。那并非少数人的责任，而是整个人类大家庭的责任。因为在人类关系中，「和平」是承认并接纳他人不可剥夺尊严的果实，也是在追求个人及民族的整体发展中，合作和承诺的果实。

值此新年伊始，我祈愿各式各样人工智能的快速发展，不致加剧普遍存在于当今世界的 不平等和不

公义情况，而是帮助这世界终止战争和冲突，减轻人类大家庭所遭受的多种 苦难。愿基督信徒、不同宗教的追随者，以及所有善心人士，在和谐中共同努力，迎接 数字革命带来的各种机会，并面对所带来的挑战，让我们的后代能有一个更精诚团结、 更富于正义与和平的世界。

教宗方济各

发自梵蒂冈

2023年 12月 8日

(台湾明爱会 恭译)

1 《論教會在現代世界》牧職憲章，33。

2 同上，57。

3 參閱：《願祢受讚頌》通諭，104，(2015 年 5 月 24 日)。

4 參閱：同上，114。

5 在密涅瓦對話 (Minerva Dialogues) 活動中向與會者致詞 (2023 年 3 月 27 日)。

6 參閱：同上。

7 參閱：〈致達佛斯世界經濟論壇會議主席的訊息〉(2018 年 1 月 12 日)。

8 參閱：《願祢受讚頌》通諭，194，(2015 年 5 月 24 日)；〈向「電子時代的公共利益研討會」與會者致詞〉，(2019 年 9 月 27 日)。

9 《福音的喜樂》勸諭，233，(2013 年 11 月 24 日)。

10 參閱：《願祢受讚頌》通諭，54，(2015 年 5 月 24 日)。

11 參閱：向宗座生命科學院大會與會者致詞，(2020 年 2 月 28 日)。

12 參閱：同上。

13 《眾位弟兄》通諭，27，（2020年10月3日）。

14 同上。

15 參閱：同上，170~175。

16 參閱：《願祢受讚頌》通諭，177，（2015年5月24日）。