

专利法常设委员会

第三十六届会议

2024年10月14日至18日，日内瓦

有关专利和新兴技术的背景文件（SCP/30/5 更新）

秘书处编拟的文件

目 录

一、导 言	3
二、人工智能技术概况：神经网络、深度学习和生成式人工智能	3
A. 机器学习	4
B. 神经网络	5
C. 深度学习	7
D. 生成式人工智能	9
E. 目前深度神经网络和生成式人工智能的局限	15
三、人工智能相关发明的专利保护	17
A. 总论	18
B. 可授权主题	20
C. 新颖性和创造性	24
D. 充分公开和权利要求	25
E. 工业实用性	28
F. 发明人权利和所有权	28
四、作为工具应用于专利制度的申请过程和管理工作的人工智能技术	30
A. 知识产权主管部门的工具	30
B. 申请人、第三方和知识产权专业人员的工具	31
产权组织和各知识产权局发布的有关人工智能的网页和出版物	32

一、导 言

1. 在 2018 年 12 月 3 日至 6 日在日内瓦召开的专利法常设委员会（SCP）第二十九届会议上，委员会商议秘书处将编拟一份有关专利和新兴技术的背景文件并提交 SCP 第三十届会议。因此，秘书处向 SCP 第三十届会议提交了文件 SCP/30/5 以供讨论。
2. 在委员会第二十九届会议的讨论中，很多发言的代表团都提到人工智能是委员会待讨论的问题。因此，尽管“新兴技术”一词可能有着宽泛的涵义，涵盖了多项不同的新技术，包括人工智能和机器学习、区块链、合成生物学及基因编辑等，但文件 SCP/30/5 涉及的是有关专利和人工智能的背景信息。
3. 在 SCP 第三十五届会议上，委员会商议秘书处将更新文件 SCP/30/5，根据从成员国和地区专利局收到的信息，汇编与人工智能相关发明可专利性有关的现行法律和实践，并提交 SCP 第三十六届会议。根据该决定，本文件对文件 SCP/30/5 第三部分 A 节至 E 节进行了更新。
4. 此外，为保持文件的完整性和可读性，秘书处在文件 SCP/30/5 第二部分中增加了对生成式人工智能的概述。此外，根据 SCP 自第三十届会议以来与专利和人工智能专题有关的活动，更新了文件 SCP/30/5 第三部分 F 节和第四部分。
5. 本文件保留了文件 SCP/30/5 的结构，由三部分组成。文件第一部分提供了关于人工智能技术的背景信息。文件第一部分阐述了人工智能技术的基本概念，尤其是构成当前人工智能发展核心的机器学习技术和生成式人工智能。这种对于技术的概述性说明被认为是必要的，因为为了解一项技术对于专利制度的影响，至少要对这项技术本身具备基本的了解。
6. 本文件第二部分和第三部分阐述了专利制度和人工智能之间的相互关系。它们处理了两个不同的问题：第二部分探讨了作为专利保护客体的人工智能技术（或人工智能相关发明），第三部分讨论了专利制度的主管部门和用户对于人工智能技术作为工具加以应用。
7. 关于“专利质量”一词，尽管无法为它找到唯一的定义，但通过 SCP 此前的活动得到了两个主要概念。它们是：(i) 专利本身（即符合实质性标准并被授予专利的发明）的质量；及(ii) 专利局和专利局以外的专利程序的质量（文件 SCP/27/4 Rev.）。从这个角度来看，可以说有关人工智能相关发明专利保护的问题涉及专利质量的第一个方面，而通过 AI 技术改进专利程序方面的问题与专利质量的第二个方面有关。
8. 此外，本文件载有一份列有世界知识产权组织（产权组织）和各知识产权局有关人工智能和专利的网页的附件。

二、人工智能技术概况：神经网络、深度学习和生成式人工智能

9. 尽管对于 AI 没有唯一的定义，但 AI 系统可被看作主要是学习系统。本文件第一部分介绍了神经网络（NN）、深度学习（DL）和生成式人工智能领域最重要的技术概念，它们是当前最为重量级的人工智能技术。¹本节使非计算机专业人士得以了解这些新兴技术如何便利地发挥作用，以帮助人们更好地了解 AI 技术和专利之间的相互关系。

¹ 2019 年世界知识产权组织技术趋势——人工智能，第 31 页。机器学习在 AI 相关专利申请中的占比为 89%，在所有 AI 相关专利中的占比为 40%。在 2013 年至 2016 年期间，深度学习在机器学习技术领域中的年增长率达 175%，神经网络增长率为 46%。

A. 机器学习

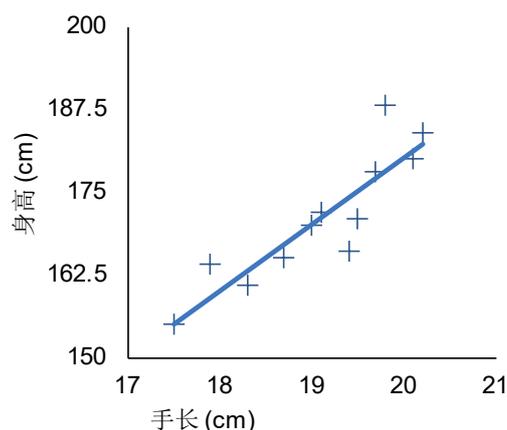
10. 从历史上来看，首批 AI 方法是对机器编程。编程在这里是指人向机器发出步进指令以完成一项任务。例如上世纪 80 年代的主要 AI 方法是专家系统，它通过各领域专家所撰写的规则复制人类的专门知识。由于这些方法成本高且具有局限性，随之而来的是发生在 1987 年至 1993 年期间的所谓第二次 AI 冬天。

11. 相比之下，机器学习（ML）方法探索的是机器如何通过输入示例和预期输出学习一项任务的解决方法，同时无需以步进指令的顺序对如何实现目标直接进行编程。这一方法更接近实际的生物性认知：儿童学习辨认物体（如杯子）的方法是通过相同物体（如各种不同种类的杯子）的示范作用。它是到目前为止最为主要和成功的 AI 方法。

12. 一般来说，机器学习法接受观察结果的输入，并把它们用于预测输出。这种学习法会试图从给出的输入输出对数据集中建立起数学模型，使它的预测和预期输出之间的差别最小化。通过这种方式，它尝试学习给定输入和输出之间的关联/模式，这些关联/模式可被代入之前未见过的新输入。

13. 为了说明这一学习过程，我们可以看一看最简单的机器学习法——线性回归。假设我们希望学习如何把人的身高和他/她的手长相互关联。已有一些身高和手长对的观察结果（左表），在下图中用十字表示：

身高 (cm)	手长 (cm)
170	19.0
155	17.5
184	20.2
188	19.8
178	19.7
172	19.1
165	18.7
180	20.1
161	18.3
171	19.5
164	17.9
166	19.4



手长 =
 $0.098 * \text{身高} + 2.23$

if 身高 = 180 cm,
手长 =
 $0.098 * 180 + 2.23$
手长 = 19.9 cm

14. 线性回归是一种在这些点之间找到误差最小的直线的方法。使误差最小化的过程就是训练。数学方法通过找到距离数据点最近的直线来实现这一训练。一旦找到了这条误差最小的线，即可根据一个人的身高预测他/她的手长。例如，如果一个人身高 180 厘米，模型会预测其手长为 19.9 厘米（见右框）。

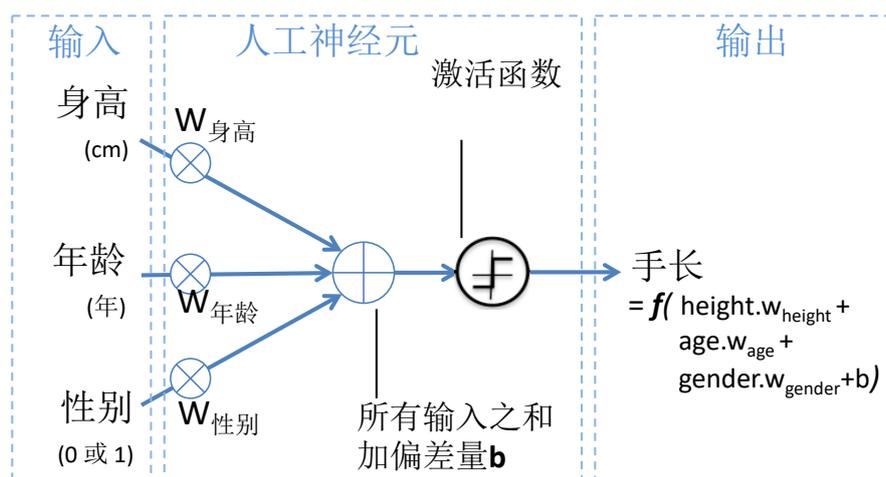
15. 这种简单的方法对于学习更复杂的问题来说无疑具有很大的局限性，例如涉及两个数值变量的问题。在上述例子中，看起来除了身高，还应考虑年龄和性别，才能使手长的预测结果更可靠。还会使用更复杂的数学模型，尤其是不仅限于直线的非线性模型。

16. 在那些更为高级的方法中，神经网络（NN）提供一种通用预测器，能够接受任何类型的输入。神经网络尤其擅长解决涉及非结构化数据作为输入（如图像或语音）的任务。作为一种高级的神经网络类型，深度学习仍然是人工智能专利申请中的核心技术。

B. 神经网络

17. 支撑神经网络的基石是人工神经元，也被称为感知器或节点。它由弗兰克·罗森布拉特在 20 世纪 50 年代和 60 年代开发。神经元接受 n 个被称为特征的输入，它们是待处理数据（像素、词语、信号等）的数值表示。每个输入乘以权重并求和（见下图）。经过加权求和的组合值与偏置量 b 相加。最后得到的数值传送至激活函数 f 。

18. 例如，回到预测手长的例子中，如果已有一个人身高、年龄和性别的数据，则人工神经元表示如下：

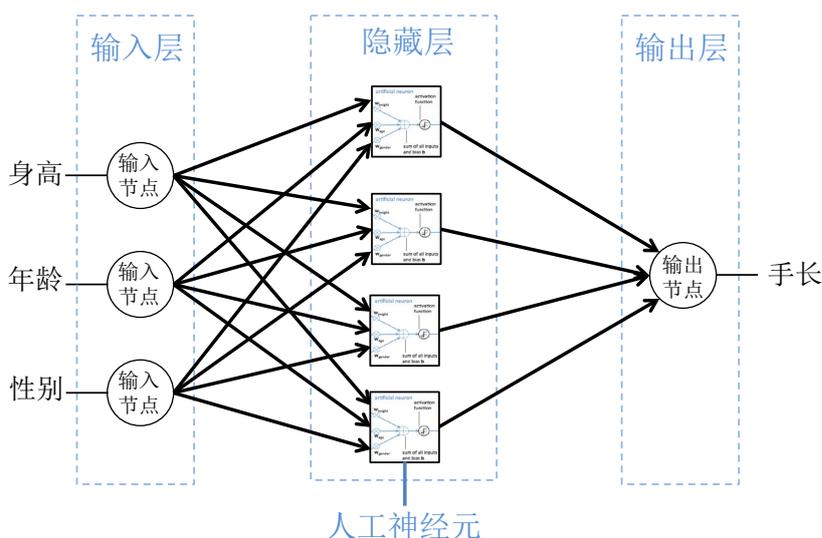


19. 权重反映了相关输入特征的强度，也就是说一项特征就其本身而言对于最终结果产生多大的影响。

20. 激活函数模拟生物神经元的“放电频率”——传播最终信号或是不传播信号。它对经加权求和的输入进行某种固定的简单数学运算。最常用的激活函数是 ReLU（修正线性单元）。²

21. 人工神经元是一个相对简单的函数。对它进行编程可以使用少于 25 行的代码。那么一个完整的神经网络至少由三层组成：一个输入层、一个或多个隐藏层和一个输出层。输入层和输出层包含不进行计算的节点。它们仅仅是为输入节点向隐藏层传递数值信息，或是为输出层把信息从网络传送到外部世界。隐藏层包含如上所述的人工神经元。来自相邻各层的节点彼此之间存在用箭头显示的连接（或边）。

² ReLU 把一个数值作为输入，得到的是 0 与该数值中最大的那个数。例如，如果输入是“1”，则输出为“1”，如果输入是“-1”，则输出为“0”。



22. 输入层充满了数字编码信息，然后通过隐藏层向前传播。初始数值被隐藏层的神经元修改，然后传播至输出层并与最终输出相对应。输出节点的数量与预期从神经网络得到的应答数量相匹配。例如在这个例子中，预期得到的是手长这个唯一值。数据在这里总是经过各层向前移动。

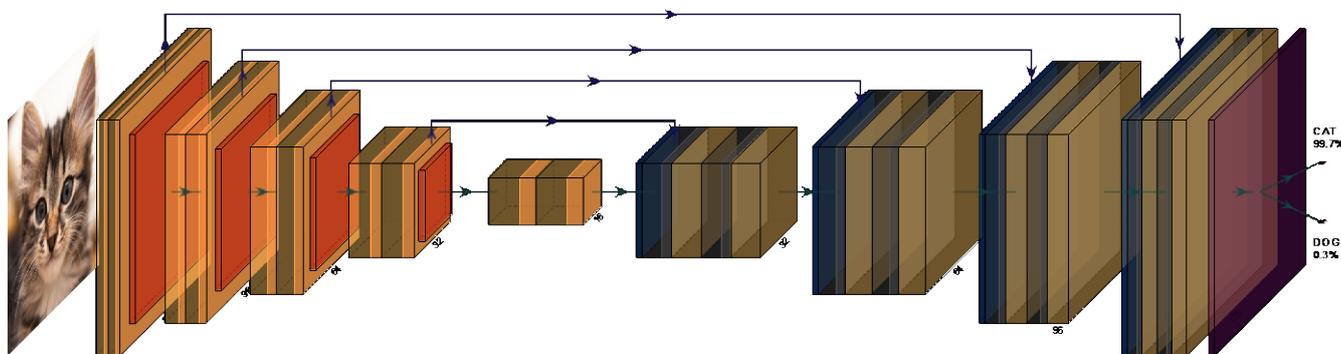
23. 训练神经网络就是对隐藏层中所有神经元的权重和偏置参数进行设置，使在一组示例中所观察到的偏差最小化，这类似于上文 A 部分中所述的线性回归。训练神经网络的机制基本上是“从错误中学习”。训练数据由一些输入/输出对组成。当神经网络有了一个输入后，它对相应的输出可能是什么进行随机“猜测”。然后它会查看其应答与实际输出之间存在多大差距，并且再对其权重和偏置量作出适当调整。所有输入/输出对都要重复这个过程，直至得到最优权重和偏置量。

24. 应注意的是人工神经元只是在很小程度上受到哺乳动物生物神经元结构的启发，并且复杂程度比后者低得多。生物神经元相比人工神经元要复杂多样得多。有很多因素（突触结构和几何结构、神经递质类型等）会影响信号传播。例如突触包含超过 2000 种不同的蛋白质，具有非常多种多样的物化特性。³

³ “人工和生物神经网络之间的区别”，Nagyfi Richárd，登于 Toward Data Science 上的博客文章，2018 年 9 月。<https://towardsdatascience.com/the-differences-between-artificial-and-biological-neural-networks-a8b46db828b7>。

C. 深度学习

25. 尽管神经网络从上世纪 50 年代开始就已为人知晓，但直到 2000 年代，隐藏层数量仍然通常仅为一个。计算能力的提升使神经网络层数在过去 10 年中得以增加（由此“深化”）。例如对于猫或狗图像分类问题（图像中是猫还是狗？），目前的深度神经网络如下所示：



26. 在以上的例子中，我们看到与前述的简单神经网络相比，目前的网络在复杂程度方面发生了很大变化：

(i) 输入节点数量很多；每个输入节点接收一个图像的一个像素的信息。在以下示例中，猫的图像被转换成一个尺寸为 18*40 的矩阵，每个像素由其灰度级定义，灰度级由 0（白色）到 256（黑色）之间的整数表示。对于写实猫狗图像的分类，我们通常使用尺寸为 128×128 像素的图像，每个像素由红绿蓝 3 个值确定，即 49,152 个输入节点，因此以下每个神经元有 49,152 个输入特征。

人对图像的感知



计算机对图像的感知

5	4	6	5	5	5	6	6	7	9	8	9	8	4	6	11	40	136	
6	8	5	5	5	5	5	5	6	7	8	12	9	4	65	27	18	61	
5	9	7	4	4	4	4	4	5	6	6	7	7	7	70	123	86	9	26
4	82	187	6	5	4	4	4	4	5	6	7	5	85	99	111	6	18	
4	103	132	77	4	3	4	4	4	5	6	10	26	31	98	10	7		
8	140	139	93	18	11	5	3	5	5	6	6	66	92	133	149	4	9	
10	149	182	78	73	6	10	4	5	5	11	8	91	120	122	160	9	16	
15	155	149	85	104	5	10	5	6	17	128	86	115	143	121	9	28		
22	147	149	142	80	101	27	18	23	7	140	85	119	137	157	129	13	46	
15	144	146	126	113	126	135	68	156	97	182	129	121	158	95	152	18	61	
18	150	149	99	135	144	217	212	176	119	190	174	112	83	94	142	36	81	
16	144	147	85	101	135	179	223	196	187	201	151	131	87	161	123	41	185	
43	110	114	115	68	93	147	187	218	197	229	165	182	136	99	98	74	117	
27	114	133	83	75	117	179	168	184	234	161	181	142	140	86	95	112		
11	84	105	121	84	153	188	183	215	171	229	198	169	161	124	56	84	189	
3	66	110	112	116	135	136	233	235	218	127	196	174	97	110	164	144	110	
2	48	84	185	113	148	149	238	183	146	74	195	187	158	173	284	168	182	
6	21	136	129	114	153	98	197	144	172	61	124	128	138	189	188	176	181	
9	11	180	148	121	122	92	244	133	287	187	70	182	117	192	147	172	132	
6	9	179	168	121	53	40	141	116	154	119	141	238	135	216	189	180	138	
8	9	165	185	190	125	242	184	133	117	138	193	284	106	223	191	193	175	
8	33	183	176	139	212	171	231	123	132	63	288	192	65	195	195	285	283	
7	12	181	233	289	127	210	188	51	164	48	197	192	184	288	217	211	216	
7	14	191	229	158	42	195	286	82	132	189	74	187	212	161	223	211	216	
6	22	282	191	185	139	181	65	170	185	136	78	187	199	174	170	287	218	
10	29	168	163	178	193	211	62	211	75	188	110	171	288	217	173	213	195	
18	21	184	137	158	158	113	193	95	156	161	134	288	184	171	286	281		
21	35	98	195	234	163	86	111	131	112	156	72	68	139	172	168	286	281	
31	46	181	161	159	155	186	151	53	197	195	93	67	114	198	167	197	284	
32	29	87	188	124	97	38	192	40	225	112	57	68	133	148	154	195	142	
40	26	66	184	194	67	36	97	48	152	127	186	119	283	178	72	185	181	
55	39	72	145	166	198	183	114	91	174	147	114	176	125	179	189	194	228	
97	72	89	173	163	185	163	91	182	114	94	98	164	156	163	114	184	224	
135	111	111	187	160	143	169	94	63	76	84	115	144	155	169	89	128	215	
163	151	115	164	184	174	126	135	74	86	98	132	158	143	123	158	187	216	
195	173	88	131	193	183	112	189	184	128	128	114	131	158	175	191	213		
199	177	178	185	158	127	112	114	181	184	112	123	167	166	163	197	187	287	
139	157	161	116	126	122	124	89	92	89	118	97	142	169	189	218	174	194	
121	112	148	122	141	118	92	98	85	88	94	91	156	214	211	152	139	185	
97	184	181	178	153	114	186	95	88	96	84	129	217	221	157	156	178	285	

(ii) 多层神经元被用于连续处理输入信息。对于图像处理，10 层以上的神经元并不鲜见，每层可能包含上百个神经元，它们通常有不同的排列方式以提供特定的优势。

(iii) 这种典型的深度神经网络可能有数千万个要在训练期间设置的权重和偏置量，需要上万个有标号图像。

27. 令人惊讶的是在像 Keras⁴这样的开源框架中，经过培训的数据科学家可以用少于 100 行的代码实施神经网络。在处理一个猫狗图像在线开放数据集的情况下，使用商用硬件的网络在分类方面的准确性可达 93%：这与人的表现相差无几（对于这类任务人的准确性估计约为 95%）。

28. 层数的递增使全球预测任务所涉及表示和过程的层级这一概念应运而生。第一批层通常捕捉输入数据中的低层级模式（就处理图像而言如线条、彩色区域等），中间层识别的是更高层级的结构（如猫狗分类情况下典型的猫耳或口鼻），最后的层自我分工，依据所识别的结构完成最终的预测任务。

29. 与传统的神经网络相比，神经网络具有若干重要特征，这也是后者在当前取得成功的原因。

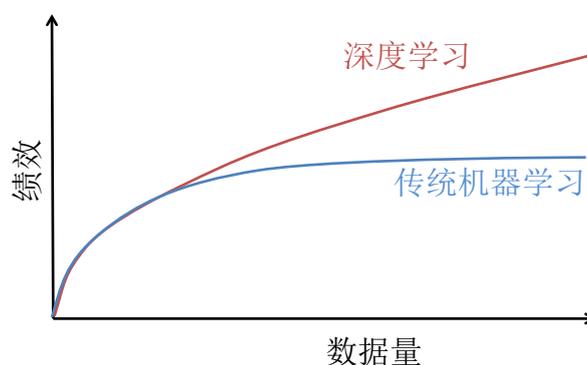
发现特征表示

30. 传统的机器学习使用工程师人工编制的特征来解决问题。例如，对于预测手长，机器学习工程师要根据自己的直觉和尝试选择一些特征，如人的身高、性别和年龄。这个步骤被称为特征工程。一个特征是机器学习算法用于预测输出的数据的一个方面。这一步普遍十分耗时，如果处理的是非结构化数据（图像、文字、语音、视频），效率还相对较低。

31. 纵观机器学习的发展史，神经网络首次显示出具有能够从原始数据中自动发现此类特征的实践能力。通过深化层数，神经网络在学习有用特征的同时还学习如何通过它们处理任务。例如在预测手长的例子中，操作人员只是向神经网络输入尽可能多的生物识别测量结果，让网络自动识别可用来作出最终遴选的数据。在图像分类的例子中也是如此，原始像素数据被发送给网络，后者对可作为判定依据的模式（如耳朵的形状、舌头或牙齿）进行识别，以判断输入的图片是狗还是猫。

数据规模和深度学习的绩效

32. 传统机器学习技术的绩效随着训练数据量的增加很快进入稳定期。这意味着增加更多的训练数据是无用功，训练算法经过一段时间后变得有些“饱和”。深度学习的重要特征之一是绩效随着训练数据的增加持续提升。这一特征解释了为什么在机器视觉中目前已有的最大网络可使用多达 1500 万个图像用于训练。



⁴ Keras: 基于 Python 的深度学习库, François Chollet 等, 2015 年-2024 年。https://keras.io/。

33. 从数学的角度来看，可把人工神经网络模型理解为仅仅是一组矩阵运算并找到导数。⁵随着计算能力的加强，深度学习可超过任何其他机器学习法，前提是有可供使用的海量训练数据。

34. 我们能在多大程度上扩展这种神经网络，并继续提高深度学习模型的绩效和能力？为探索和提高这种扩展能力而开展的实验使生成式人工智能领域在当前取得了巨大成功。

D. 生成式人工智能

35. 生成式人工智能可以定义为一种能够在用户的提示下创建新的、有意义的、类似人类的内容，包括文本、图像、音频或视频。“提示”此处指的是通常由人类用户生成的自然语言文本指令。

36. 下一节将介绍生成式人工智能的主要概念：

- 判别任务和模型与生成任务和模型之间的区别；
- transformer 架构和注意力机制，它们是现代生成式人工智能兴起的基础；及
- 大型语言模型的主要特点。

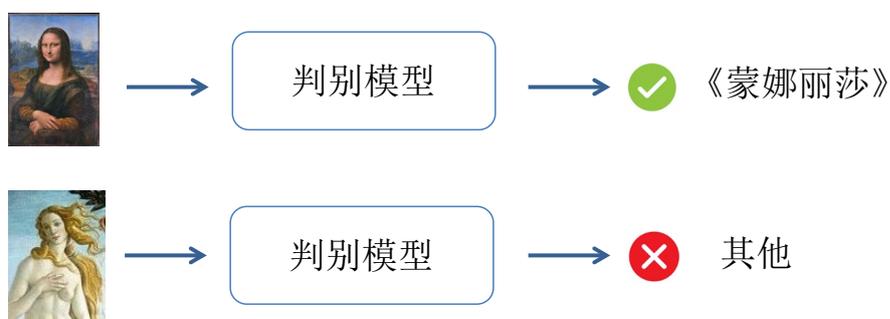
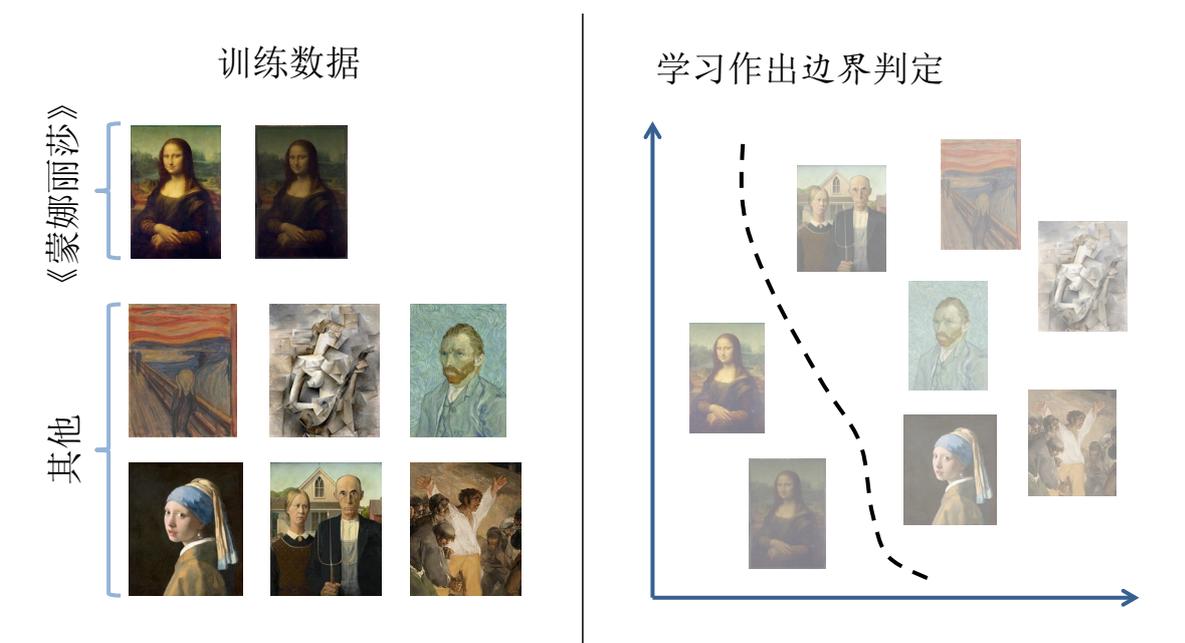
判别任务与生成任务

37. 神经网络通常可适用于两种不同的任务：

- 判别任务是对输入数据进行判定的任务，如分类、识别文本中的名称或分割图像。判别模型是为将输入数据分为这些不同类别而调整和训练的模型。
- 生成任务是指在给定输入数据的情况下创建新数据的任务。生成模型是为创建此类新数据而调整和训练的模型。它们通常用于翻译文本、生成图像、总结文本或回答问题。

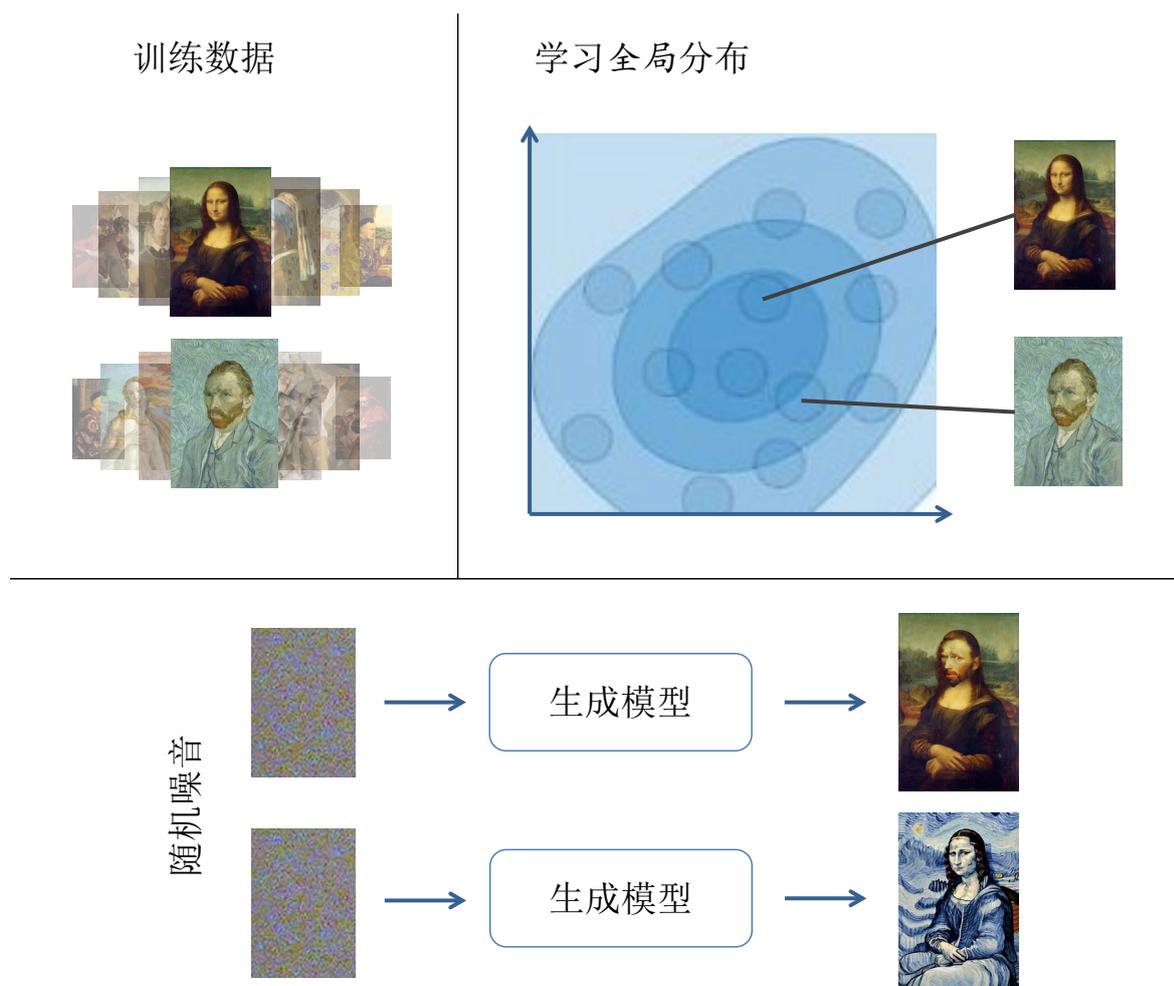
38. 以下两张图说明了机器学习任务的这两种基本类型。

⁵ 对于向量处理器，可对这种数学计算进行高度优化（针对大量数据点多次重复进行相同计算），以及通过 GPU（图形处理器，与用于加速电子游戏的装置相同）或新的专门硬件把计算速度提高几个数量级。



来源：绘画图形来自公有领域下的维基共享资源（Wikimedia Commons），原图来自 2024 年产权组织《生成式人工智能》专利态势报告。<https://doi.org/10.34667/tind.49740>

39. 上图是对一幅图像是否为绘画作品《蒙娜丽莎》进行分类的判别任务。该模型必须学会如何区分两类图像：《蒙娜丽莎》和其他绘画作品。对于这类模型，学习过程主要集中在区分类别的标准上。因此，如果以绘画特征的空间来表示模型，该模型重点在于表示两类绘画之间的边界。



来源：绘画图形来自公有领域下的维基共享资源（Wikimedia Commons），原图来自 2024 年产权组织《生成式人工智能》专利态势报告。<https://doi.org/10.34667/tind.49740>

40. 上图是生成新绘画样本的生成任务。在生成任务中，模型必须学习每幅绘画的全局特征，以便生成连贯的新绘画。对于这类模型，学习的重点是表示绘画特征的全局分布。生成这两幅图像使用的是原版 Stable Diffusion⁶模型。

41. 判别模型擅长分类，但不能生成新数据。相比之下，生成模型也能处理判别任务，但准确率通常低于判别模型。与判别模型相比，生成模型的参数更多，计算成本更高，需要的训练数据也更多。因此，判别任务最初更受关注。随着算力的提高和可用的训练数据越来越多，生成任务变得更加可行，并在过去几年吸引了更多的开发活动。

Transformer 架构

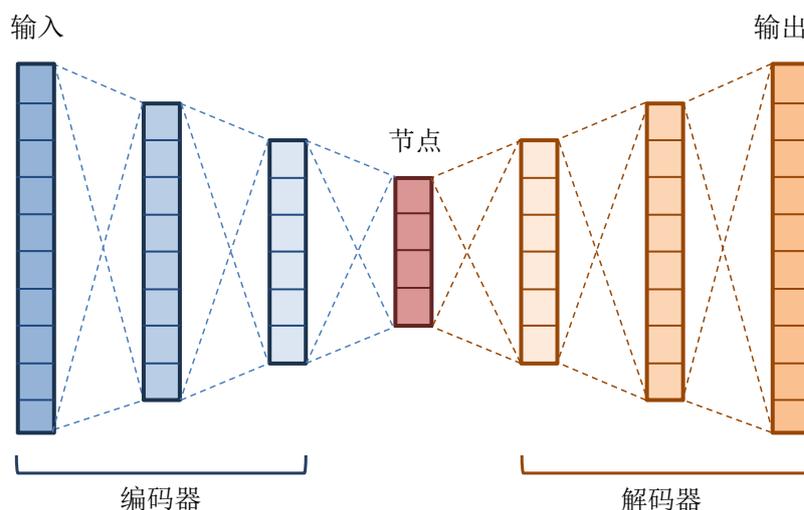
42. Transformer 是一种专为自然语言处理（NLP）任务设计的神经网络。它于 2017 年首次被提出⁷，与其他传统深度学习方法相比，它经过开发在训练过程中具有更好的扩展性。Transformer 基于两个重要概念：编码器-解码器和自注意力。

⁶ Rombach R.、Andreas Blattmann A.、Dominik Lorenz D.、Patrick Esser P.、Björn Ommer B. (2021)。《使用潜在扩散模型进行高分辨率图像合成》。arXiv:2112.10752, <https://arxiv.org/abs/2112.10752>。

⁷ Vaswani A.、Shazeer N.、Parmar N.、Uszkoreit J.、Jones L.、Gomez A.、Kaiser L.、Polosukhin I. (2017)。《注意力就是你所需要的一切》。神经信息处理系统进展 30。

43. **编码器-解码器**模型基于三个部分：编码器、代码和解码器。编码器是一个神经网络，它学习如何将输入数据编码并压缩成中间表示形式，也就是基本上为一串数字的代码。然后该代码在解码器中使用，解码器是另一个神经网络，它学习如何解压缩数据并将其重构为预期的输入格式。除了数据压缩，编码器-解码器的目标还在于学习如何表示某些数据的性质，以便对这种内部表示法进行修改后仍能重新构建出有意义的新输出。

44. 这种方法尤其针对机器翻译而开发，在机器翻译中，源语言的输入被编码为压缩的数字表示，然后根据大量的翻译输入/输出示例将其解码为目标语言。



45. 上图是编码器-解码器神经网络架构。输入（如文本或图像）通过若干层神经网络编码成向量。编码是一种中间压缩数字表示法，解码器可使用编码以所需的目标语言或图像表示法生成输出。需要注意的是，现代模型中的层数可能比上图显示的要多得多。

46. 对于专门用于判别任务的系统，通常会移除解码器部分，因为一般来说只输出分类，而不是新创建的数据。BERT（Transformer 双向编码器表示）模型系列就是这种情况，它是 2018 年至 2022 年期间机器学习领域的关注焦点。对于专门用于生成任务的系统，编码器可能会被移除，因为输出应该是新生成的数据，而编码器可能没有帮助。OpenAI GPT（生成式预训练 transformer）模型就是这种情况。

47. **注意力**是一种神经网络机制，能让学习过程同时关注文本的不同部分。它可以捕捉文本中的长距离依赖关系，这对于理解复杂的语言并对其进行表述非常重要。在实践中，**注意力**会根据重要性为输入词分配权重，从而使模型更加重视相关词。然而，如何选择这些重要的词呢？根据这一选择标准，存在着不同类型的注意力机制。

48. 最初的 transformer 架构采用的是**自注意力**机制：根据这种方法，会对训练数据所有输入中每一对可能出现的词之间的相关性进行评估。让我们快速了解一下这个过程：如果两个词在训练数据中出现的频率比随机出现的频率更高，那么它们的权重就会比不相关的词更高。

49. 总之，**Transformer** 是一种编码器-解码器深度神经网络，它利用注意力的优势，随着层数的增加，其训练时间比以前的深度神经网络更快。扩展能力与层数直接相关。堆叠更多的层意味着神经元全局数量成倍增加，从而可以存储更多的学习参数。在 Transformers 中，每一层都使用注意力。因此，学习过程不会对输入进行多次循环，这种循环会造成很高的计算成本。相反，Transformer 会让输入经过多个注意力层。

50. 如下表所示，随着时间的推移，Transformer 的层数不断增加，从而可以在更大的数据集上进行训练，学习更多的参数信息。

模型名称	开发者	年份	层数	参数数量	训练数据量（估计数）
Transformer (原版)	谷歌	2017	6	1.1 亿	8 亿单词
BERT	谷歌	2018	12	1.1 亿	33 亿单词
GPT	OpenAI	2018	12	1.1 亿	6000 亿单词（40GB）
GPT-2	OpenAI	2019	48	15 亿	8 百万网页
GPT-3	OpenAI	2020	96	1750 亿美元	570GB 文本数据
GPT-4	OpenAI	2023	120	1.8 万亿（估计数）	13 万亿词

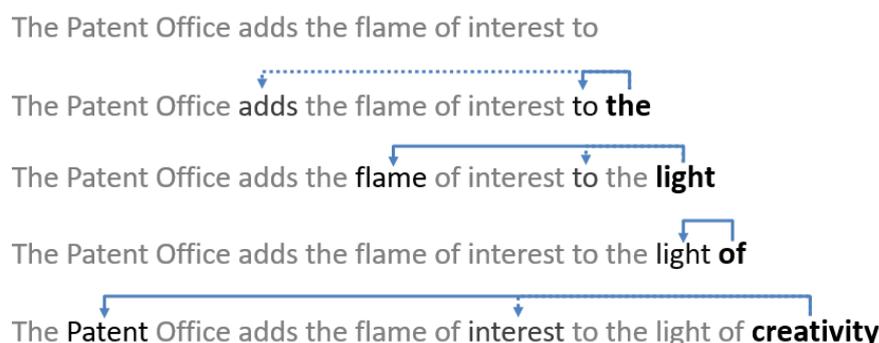
51. 我们之前提到，模型的生成能力会随着模型规模和训练的增加而提高。在这里，我们明白了为什么 transformer 架构使生成式人工智能得以兴起，而判别任务却位列最成功的人工智能系统之前。

从大型语言模型到 ChatGPT

52. 如上表所示，大型语言模型（LLM）是使用海量训练数据训练出来的超大模型。它们的开发需要大型现代化计算基础设施。即使是拥有十万个处理器的大型计算中心，此类 LLM 的训练通常也需要两周到三个月的时间。这些模型是众所周知的 OpenAI ChatGPT、微软 Copilot 或谷歌 Gemini 等现代助手的基础。它们共享一种称为自回归的特殊生成机制。

53. **自回归模型**不是一次性生成完整的文本或图像，而是根据之前生成的要素，一次生成一个要素的数据。例如，对于文本，生成过程是一个迭代过程的结果，即根据之前生成的单词预测下一个单词，直到预测出文本的结尾。这可以看作是模仿人类语言的自动完成功能。这也与众所周知的智能手机键盘的下个词预测功能类似。不过，这里所考虑的模型要比手机应用程序的下个词预测大得多，通常是后者的 1000 倍或更多。

54. 这种方法的一个主要优势是，对下个词或字的预测可以通过普通文本来学习。由于在训练过程中会扫描文本以设置模型的内部参数，因此这些模型有能力在给定文本上下文的情况下预测最有可能出现的下个词。这种训练不需要人工标注数据。计算机可以在没有人工指导的情况下学习识别模式，这种学习方式被称为无监督学习。然而，要达到这些模型目前的绩效，需要海量文本。



55. 上图是通过自回归生成的文本示例，说明了什么是注意力机制。当模型逐字生成文本时（粗体字），它可以通过注意力机制（蓝色箭头）关注与每个新生成的单词相关的先前单词。在实践中，可以达到的字词窗口规模非常大，最新的模型可以达到 100,000 个词。

56. 这种模型有两个有趣的特性：

- 它们不仅要学习一般语言，还要学习如何在训练数据中生成与世界上的实体和事件有关的各种事实的文本。
- 它们能够记住如此之多的文本内容，以至于如果提供足够的初始上下文，完成的内容就会变得出奇地精确和相关。

57. 这两个特性使得像 ChatGPT 这样的多功能助手成为可能，它们不仅能生成语法正确的文本。这些助手还可以回答问题，并实现各种任务，如总结、分类、改写、翻译、创作故事等。然而，考虑到这些模型当初的设计初衷只是为了预测下个单词以完成输入文本，它们如何能完成如此复杂的任务呢？

58. 这种协助用户和与用户对话的能力是通过用成千上万的**指令和预期结果**示例来调整这些模型来实现的。这些指令与用户为表达其要求而输入的“提示”相对应。模型将学习生成与预期结果相匹配的文本，以完成这些指令。基础 LLM 具备生成大量不同人类语言文本的能力，而这种额外的调整则训练模型学会如何使用这种能力来按照指令执行任务，从而实现流畅而多变的对话。

59. 以往传统的“监督式”机器学习模型需要大量特定任务注释的训练数据。相比之下，这些 LLM 只需写入自然语言“提示”即可生成新内容。因此，使用这些生成式人工智能工具不再需要技术技能。普通大众可以直接使用现代尖端人工智能，只需用自然语言指令提示模型即可。

60. 目前已开展了大量研究工作，以处理多个业务领域的各种任务和协助，以及生成其他语言形式，如编程代码或描述蛋白质的结构。有关应用和实例的综述，请参阅最近发布的产权组织生成式人工智能专利态势报告。⁸

专为生成发明定制的 LLM

61. 对于 SCP 来说，最相关的应用之一是**人工智能生成的发明**。为了实现这种能力，通常会对 LLM 进行微调，加入不同或额外的指令以及与发明相关的预期结果。此外，还可以利用专利文件、科学文本或其他技术内容对 LLM 进行进一步的预训练，并可能结合索引此类文献的搜索引擎（一种称为“检索增强生成”的技术）。

62. 这种定制模型可以有不同的用途：

- (i) 模型用于帮助撰写专利申请，以一项人类发明的要点为基础，包括相关现有技术的建议。
- (ii) 人类识别发明，并利用模型来识别这些发明可能解决的其他问题。
- (iii) 人类识别问题，并利用模型生成可能的解决方案。
- (iv) 人类利用模型识别潜在的有价值问题，并为这些问题生成可能的解决方案。

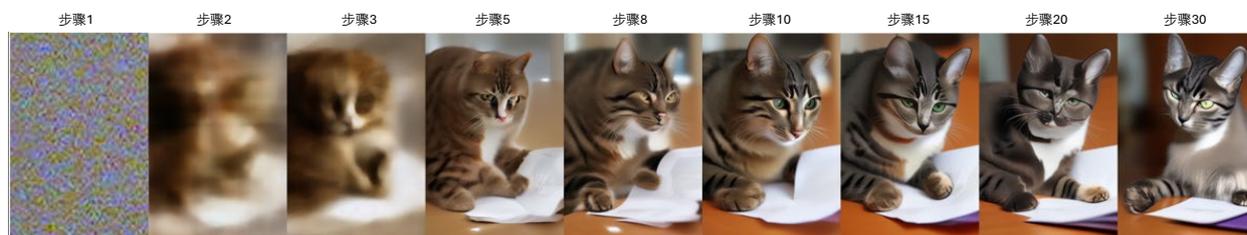
⁸ 世界知识产权组织（产权组织）（2024 年）。《生成式人工智能》。《专利态势报告》。日内瓦：产权组织。
<https://doi.org/10.34667/tind.49740>。

63. 定制化 LLM 的这些不同用途可被视为逐步提高的发明人权利，从为人类发明人提供协助和见解，到更多地生成新构思。然而，还有其他因素决定了人类在生成创造性 LLM 输出方面所作工作的性质和程度：

- (i) 创建用于查询模型的提示；
- (ii) 为指导模型而提供的技术原则和特征；
- (iii) 模型参数的选择；
- (iv) 提示 LLM 的迭代次数。

图像和其他模式的 LLM

64. 我们在本文件中着重讨论文本生成。然而，在 LLM 之前就出现了成熟的图像生成模型，它们使用的是不同类型的模型。有意义的新图像的构建依据的是对十万张图像进行类似训练的模型，并由额外的机器学习机制进行控制。对训练图像的文本描述使文本和图像模式的结合成为可能。例如，扩散模型可以根据简短的文本描述生成高分辨率图像，2022 年发布的著名的 Stable Diffusion 模型就是一例。



来源：Stable Diffusion 模型，提示为“一只阅读专利的猫”，原图来自 2024 年产权组织《生成式人工智能》专利态势报告。<https://doi.org/10.34667/tind.49740>

65. 上图中，图像生成使用的是原版 Stable Diffusion 模型，提示为“一只阅读专利的猫”。生成过程相当于：(i) 提供带有随机噪音（随机像素）的图像，(ii) 反复使用神经网络去除噪音。提供的文本提示用于驱动生成。随着噪音被逐步去除，利用通过大量训练数据学习到的与提示统计相关的图形元素就会构建出一个有意义的新图像。此类模型成功覆盖的其他模式包括视频（将一个图像连续转换为另一个图像）、语音、音乐以及由一幅或多幅二维图片构建的三维场景。

E. 目前深度神经网络和生成式人工智能的局限

深度神经网络是黑匣子

66. 与更为传统的算法相反，神经网络在学习过程中捕捉到的决策过程无法以人可以理解的方式直接表达出来。如上所述，深度神经网络可以自己学习数据中有用的特征。例如在猫狗分类的例子中，网络可以识别典型的猫耳或口鼻。但在实践中，这些特征大多数情况下无法被人解读。这些模式出自隐藏层中的数值优化过程，我们无法对其进行解读。

67. 此外，不可能以标准数学的形式显示对输入和输出之间关系进行定义的等式或系数。网络是最终的关系等式，可能涉及上千亿个参数。如此复杂的决策过程无法通过流程图或任何传统的算法表示法来说明。这就解释了为什么人们常常说深度神经网络是“终极”黑匣子，缺乏透明度。训练本身是由神经网络独自实现的，并且由此产生的网络极为复杂。

深度学习需要很多数据

68. 一个令人惊讶的观察结果是，神经网络和深度学习就所涉及的数学建模而言属于最简单的机器学习模型。常见的说法是一名优秀的高中生就能理解其中的基本数学原理。但它们仍提供了到目前为止最好的结果。原因是它们对于利用很大规模的训练数据集最为适应。深度学习和生成式人工智能在当今的成功与其说是由于理论发展，不如说是由于计算能力的显著提升和可用的海量人类行为数据：我们通常称之为蛮力。

69. 深度学习明显存在的局限是蛮力不可能发生的情况。这其中尤其包括没有数据或数据有限的任务（如处理少见的人类语言、发现治疗罕见病的药物等），或是任务所在领域存在法律限制。

现实世界的数据存在偏差

70. 深度学习的成功取决于有大量的可用数据，但这种对于大量数据集的依赖也产生了若干问题：

- 数据偏差：大规模的数据采集通常无法保证中立，一些群体在年龄、性别和民族方面的代表人数可能不足或过多。⁹偏差可能来自于数据采集法、已有的社会偏见或是创建数据集和模型的人员并非来自多元化群体。
- 偏差放大：机器学习训练法天然地趋向于识别数据中的判别模式以快速提升预测或生成质量。因此，它们不仅学习我们实际存在的偏差，还经常将其放大。
- 缺乏可再现性：由于模型依赖于专门的训练数据组成，再现一些所要求的结果只在非常罕见的开放数据的情况下才有可能发生。

可信性

71. LLM 以令人信服的方式生成不准确内容的情况并不少见。这种不真实的输出被称为幻觉。尽管近一段时间以来专门开展了很多工作以降低此类事实错误的发生频率，但这类现象仍然很严重，尤其是在专业领域。

72. 最近的一项研究估计，在简单任务中，GPT-4.0 每次输出产生幻觉的频率约为 28.6%¹⁰，这可能高于用户通常的认知水平。例如，当将人工智能应用于法律领域时，斯坦福大学研究人员的一项研究¹¹测定，GPT-4.0 每次响应产生幻觉的频率至少为 58%。根据同一组研究人员的一项研究，即使是专门使用法律文本进行预训练并使用法律数据库来丰富和控制提示（一种称为“检索增强生成”或“RAG”的技术）的定制化 LLM，如律商联讯和汤森路透的商业化 LLM，其每次生成响应的幻觉率估计仍在 17%到 33%之间。¹²

⁹ 亚马逊弃用对女性表现出偏见的秘密 AI 招聘工具，Jeffrey Dastin. 路透社商业新闻，2018 年 10 月 (<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>)。

¹⁰ Chelli M、Descamps J、Lavoué V、Trojani C、Azar M、Deckert M、Raynier JL、Clowez G、Boileau P、Ruetsch-Chelli C. ChatGPT 和 Bard 的幻觉率和参考准确性系统回顾：比较分析。J Med Internet Res. 2024 年 5 月 22 日；26:e53164. doi: 10.2196/53164。

¹¹ Matthew Dahl、Varun Magesh、Mirac Suzgun、Daniel E Ho, 《大型法律拟制：剖析大型语言模型中的法律幻觉》，《法律分析期刊》，第 16 卷第 1 期，2024 年，第 64-93 页，<https://doi.org/10.1093/jla/laae003>。

¹² Varun Magesh、Faiz Surani、Matthew Dahl、Mirac Suzgun、Christopher D. Manning 和 Daniel E. Ho, 《免于幻觉？评估领先人工智能法律研究工具的可靠性》，斯坦福大学，2024 年即将出版。

73. 对 LLM 进行长期评估似乎很困难。在经典的机器学习评估中，系统不应通过其训练数据进行测试和评估。这一原则确保了不同的系统可以使用开放访问的数据集，以公平的方式进行长期基准测试和比较。LLM 是在互联网的大量数据上进行训练的，这意味着 LLM 可能会使用这些现有的、通常在网上发布的评估数据集进行训练，而且可能是多次训练。这种现象被称为数据污染。这样，对 LLM 模型的评估就不再是衡量其正确回答未见问题的能力，而是简单地记忆已发布的解决方案的能力。

74. 一些研究人员还描述了未来与可用训练数据质量相关的绩效逐渐下降的情况，这种情况被称为“模型崩溃”。随着生成式人工智能系统的输出充斥互联网，未来的模型本身可能会通过早期模型创建的在线内容进行训练，其中包括早期模型的所有偏差和错误，从而导致准确性和能力的逐步下降。

冒犯性和敏感性内容

75. 生成式人工智能模型通过海量数据、文本和图像进行训练。不可能从预训练数据中完美地过滤掉所有不恰当的、在道德伦理方面具有挑战性的内容。因此，这些模型可能会在生成的输出中复制和重新组合类似的冒犯性内容。

76. 有多种技术可以避免这种情况的发生。例如，预测的单词或短语会不断经过所谓的“护栏”进行过滤，以去除冒犯性内容（例如，“护栏”可以是黑名单上的单词）。不过，目前还不能保证完全消除此类生成输出，尤其是在使用有毒提示的情况下。

人工智能资源和基础设施

77. LLM 预训练成本高昂，而大规模利用 LLM 同样需要大型计算中心和电力。生成式人工智能也依赖于互联网规模的数据。结合投资能力和当地数字技能的可用性，多个论坛都在讨论高收入国家与中低收入国家之间的差距问题。¹³与此同时，人工智能技术也被用于改善发展中国家的教育、医疗保健和金融普惠等方面。¹⁴

三、人工智能相关发明的专利保护

78. 本文件的这个部分探讨了人工智能相关发明的专利保护。“人工智能相关发明”可能有不同的形式。创新可能在完善人工智能技术的过程中出现，也可能体现在把人工智能技术在已有设备中集成以提升其功能性或增加新功能。此外，人工智能技术可作为一种研发工具用于创造新发明。对于不同形式的人工智能相关发明，人工智能技术对于专利法不一定产生相同的影响。

¹³ 例如，见联合国和国际劳工组织的报告《注意人工智能鸿沟：塑造未来工作的全球视角》，2024年7月26日，ISBN：9789211066524，<https://www.ilo.org/media/581631/download>。一份研究论文指出，收入水平、青年人口比例、数字基础设施、高技能可交易服务专业化、英语熟练程度和人力资本与更多的生成式人工智能采用率密切相关。见《究竟谁在使用生成式人工智能？》，政策研究工作文件 10870，世界银行集团：<https://documents1.worldbank.org/curated/en/099720008192430535/pdf/IDU15f321eb5148701472d1a88813ab677be07b0.pdf>。

¹⁴ 例如，见《倾斜的天平：人工智能对发展中国家的双重影响》，世界银行博客：<https://blogs.worldbank.org/en/digital-development/tipping-the-scales--ai-s-dual-impact-on-developing-nations>。

A. 总论

79. 专利制度应为了全社会的利益，通过给予技术生产者和技术知识使用者平衡的权利和义务，为推动技术创新以及为技术转让和传播作出贡献，这一点已得到广泛认可。各国为此提供法律框架，并颁布法律法规，法院对这些法律法规进行解释，辅以行政机关所制定的实务指导。

80. 由于专利制度对于技术持中立的态度，因此每当有新技术出现时，经常会提出一个问题，即专利制度能否继续达到其目的。半导体技术、计算机软件、信息技术和生物技术皆是如此：随着技术的发展，辩论仍在继续。因此人工智能的出现引发了同样的问题和辩论，这并不足为奇，这些问题和辩论仔细研究了目前的专利制度是否已为兼顾人工智能技术做好准备。

81. 近数十年来，包括硬件和软件在内的计算机技术被应用于协助人类在很多技术领域进行发明创造。例如，计算机辅助设计（CAD）助力了机械和电子行业的发展，生物信息学为研究人员分析和解读生物数据提供了便利，以及计算化学帮助化学家找到新化学物质。计算机还被集成进设备和装置以实现某种功能。

82. 对于计算机技术来说，与其相关的新发明可归为三类：

- (i) 改进计算机本身的计算功能的新发明；
- (ii) 为实现某种功能集成了计算机的新发明（设备、装置等）；及
- (iii) 在计算机的帮助下创造的可属于任何技术领域的新发明。

83. 对于人工智能技术也可以作出类似的分类：

- (i) 涉及人工智能核心技术本身的新发明；
- (ii) 集成了人工智能技术的新发明（如集成了人工智能深度学习的翻译设备，以及用于诊断某种疾病的医疗设备）；及
- (iii) 在人工智能技术的帮助下创造的新发明（如人工智能技术帮助找到的一种新材料）。

84. 在人工智能技术现阶段的发展中，人类的指示和干预仍是创造这些发明过程中的重要部分。但是，随着人工智能技术的发展，人类干预在创造过程中的性质相比人工智能系统更多的自主执行可能会发生变化。

85. 因此，可以从另一个角度理解人工智能相关发明，着重于核心发明概念的创造。从这一角度来看，人工智能相关发明可分类如下：

- (i) 由人类识别问题和构想解决方案，而人工智能技术仅应用于对人类的解决方案进行验证、自动化、适应或泛化；
- (ii) 由人类识别问题，而解决方案的构想由人工智能技术辅助、指导或主导；
- (iii) 由人类和人工智能技术共同识别问题和构思解决方案；及
- (iv) 由人工智能技术识别问题和构想解决方案而没有任何人类干预。

在第二种情境中，人工智能技术在发明创造过程中的相关性可以从最小化到决定性。第四种情境，即通用人工智能或超级人工智能¹⁵，是当前技术尚无法达到的阶段。¹⁶但是，这一发展的可能

¹⁵ 这意味着 AI 系统能够成功执行人类大脑可完成的任何智力任务，或者机器的假设能力远远超过人类大脑。

¹⁶ 2019 年世界知识产权组织技术趋势——人工智能，第 19 页。

性与传统的计算机技术之间存在显著差别。这种差别意味着就人工智能专利活动而言，人们需要提出不同性质的新问题。

86. 自人工智能技术出现以来，创新和研究人员就开始为此类发明提交专利申请并被授予专利。如在《2019 年世界知识产权组织技术趋势——人工智能》中所述，它们涉及在一系列不同人工智能申请领域¹⁷中的多项不同人工智能技术¹⁸和多种人工智能功能应用。¹⁹开源（或开放式创新）的方法在人工智能开发人员中也很流行。²⁰关于人工智能相关发明的专利态势详细数据，请参阅所述产权组织出版物。随着人工智能在公众中的地位不断提高，科学文献中对人工智能的关注度不断增加，近年来与生成式人工智能相关的专利也呈指数级增长。²¹从专利申请量来看，最重要的生成式人工智能模型是：(i)生成对抗网络（GAN）；(ii)变分自编码器（VAE）；(iii)基于解码器的大型语言模型（LLM）。这些专利并不集中在某一特定领域或行业，而是涉及大量关键应用领域。²²

87. 关于人工智能技术如何影响专利法的问题尚无定论。许多国家尚未制定审查人工智能相关申请的特别程序。²³但人工智能技术的一些特性似乎指向了专利法中可能在未来，如果不是眼下立即，受到这一新兴技术影响的领域。可供考虑的方面包括如下几点：

(i) 由于人工智能技术主要由软件实施，目前有关计算机实施发明和使用软件的发明的专利法问题对于人工智能技术可能仍然具有相关性；

(ii) 人工智能技术的认知性需要我们进一步考虑如何把这项技术融入人类创新进程，以及它对于专利制度和专利法所作出的“人造”发明假设的影响；

(iii) 在完全再现和说明在深度学习神经网络中所采用方法方面存在的固有的技术局限性使我们注意到，它们对于专利制度基本原则之一——传播新技术知识——具有潜在影响。

88. 只要专利制度的出发点仍是为推动技术创新以及转移和传播技术作出贡献，专利制度就要继续为人工智能领域也提供激励创新的举措和分享新知识的机制（除非有其他能够充分处理这些事项的法律/社会/经济工具）。在政策层面，主要的考量可以是：考虑到专利制度的目标，人工智能技术的发展是否会破坏专利制度所力求实现的平衡？如果是这样，如何恢复这种平衡？根据人工智能技术的发展对专利法和实践进行更新是否有帮助？专利制度的现有法律概念和人工智能的兴起之间是否存在，或将存在任何缺口？

89. 为了回答这些问题，需要了解人工智能相比传统的计算机技术具有哪些技术特性，以及需要对当前的法律和实践在目前和今后适用于人工智能技术的可能方式进行评估。本背景文件并非希望全面阐释所有问题。但以下段落提供了一些专利法问题的示例，它们在为人工智能相关发明寻求专利保护和授予专利方面可能具有相关性。“人工智能相关发明”一词是指上文第 83 段和第 85 段所提及的各种不同的发明。在这个阶段，几乎没有专门处理适用于人工智能相关发明的专利法问题的正式指南。由于人工智能是一项新技术，判例法尚未得到充分发展，有为数不多的专利

¹⁷ 例如交通运输、电信以及生命科学和医学。

¹⁸ 例如机器学习、模糊逻辑和逻辑编程。

¹⁹ 例如计算机视觉、自然语言处理和语音处理。

²⁰ 2019 年世界知识产权组织技术趋势——人工智能，第 109 页。

²¹ 世界知识产权组织（产权组织）（2024 年）。《生成式人工智能》。《专利态势报告》。日内瓦：产权组织。
<https://doi.org/10.34667/tind.49740>

²² 例如，软件、生命科学、出版、商业解决方案、工业、交通、安全和电信。

²³ 例如俄罗斯联邦，见俄罗斯联邦答复 C.9199 的评论意见。

局发布了指南以厘清它们在该领域的做法。随着人工智能集成产品在市场中的进一步商业化发展，权利要求解释背景下的人工智能专利维权和许可也可能成为未来讨论项的一部分。一般来说，对许可协议进行谈判和解决专利争议需要复杂全面的考量。人工智能相关发明本身是否会使这一原本就很复杂的问题进一步复杂化仍有待观察。

90. 目前的专利制度建立在这样的假设基础之上，即某些激励机制会促进人类的创造活动。从高级别政策的角度来看，人工智能技术发展所展示出的潜力对专利制度的激励理论构成了一个法律哲学问题。人工智能机器能够以自主的方式全面处理各种不同数据（不仅是科技数据，还有个人和行为数据以及社会和法律数据）、识别问题、通过新发明解决问题以及为市场生产新产品以满足人类需求，尽管这还只是科一种遐想，但一旦实现，上述问题将尤为突出。尽管这可能是一个在智力上耐人寻味的问题，但它远超出了本文件的范围。

91. 人工智能的最新进展，尤其是在深度学习和生成模型等领域的进展，使得人工智能系统越来越先进，能够在极少人工干预的情况下完成复杂的任务。这些发展引发了有关人工智能生成发明对现行专利制度影响的重要讨论。²⁴在现行专利法下人工智能能否被视为发明人的争论愈发引人关注，一些司法管辖区还就此发布了指导意见。

92. 此外，人工智能与各行各业的快速融合也凸显出为了专利申请过程中的法律确定性、一致性和公平性，有必要就人工智能相关发明的可专利性制定更明确的指南和标准。人工智能能够自主学习和演化这一特性对新颖性、创造性和充分公开等传统可专利性要求提出了挑战。一些人工智能模型的黑匣子性质，尤其是在深度学习方面，使得发明说明书必须清晰及可再现的要求变得更为复杂。

B. 可授权主题

93. 一般来说，所有技术领域的任何发明，无论是产品还是方法，都可授予专利，条件是发明符合所有法律要求，包括发明不属于不可授权主题的要求。对于“发明”一词在国际层面没有强制性定义，国内法对不可授权主题的范围进行界定，使其与该国是缔约方的国际条约相一致。因此，各国对于可授权主题范围的界定有所不同。²⁵很多国家把数学方法、机制、从事智力活动的规则和方法、商业规则和方法以及计算机程序排除在可授权主题之外。其中一些国家作出了澄清，即这些主题仅在专利申请与该主题本身相关的范围内被排除在可授权主题之外。在一个管辖区²⁶，判例法规定，针对自然规律、自然现象和抽象概念的权利要求排除在专利保护之外。在另一个管辖区，其专利法²⁷规定“发明”是“技术概念利用自然规律实现的高度先进的创造”，并且产品发明的类别中包括计算机程序和由等同于计算机程序的电子计算机处理的任何其他信息。

²⁸

94. 除了对运行人工智能功能的硬件组件所作的改进外，与人工智能技术和人工智能功能应用有关的发明主要指向软件。因此，许多成员国认为人工智能相关发明是计算机实施发明的一个子

²⁴ 例如，见澳大利亚和巴西答复 C. 9199 的评论意见。

²⁵ 见“国家/地区专利法的若干方面——从可授权主题中排除的主题”：
https://www.wipo.int/scp/en/annex_ii.html。

²⁶ 美利坚合众国。

²⁷ 日本专利法 2(1) 和(4) 条。

²⁸ 更多有关从可授权主题中排除的主题和计算机实施发明的专利适格性的信息可参见 SCP/13/3 和 SCP/15/3（关于计算机程序不属于可授权主题的内容请特别参见 SCP/15/3 附件二）。

集。²⁹与传统计算机技术的情况一样，人工智能应用也可用于非技术领域，如金融、保险、电子商务等。此外，机器学习基于计算模型和算法来分类、聚类、回归和降维，这可能被视作数学方法。另外，尽管不能否认训练数据对于机器学习绩效的重要性，但只是属于信息的数据本身不是可授予专利权的发明。

95. 计算机实施发明或软件实施发明的专利适格性已经是难以在适格主题和不适格主题之间作出明确区分的领域之一。例如在很多国家，要求保护发明的“技术性”被认为是判断专利适格性的重要因素。在这些国家建立了判例法和主管局实践，以厘清“技术问题”、“技术手段”、“技术效果”和“技术目的”等概念。至于适用于人工智能相关发明的专利适格性要求，一些专利局发布了与人工智能相关发明有关的指南，具体如下。

96. 在澳大利亚，澳大利亚知识产权局认为，就主题适格性而言，人工智能相关发明通常是计算机实施发明的一个子集。澳大利亚法院尚未审理过包括或使用人工智能的发明，但一般认为只要为技术问题提供了某种技术解决方案，就具有可专利性。根据这些原则，如果人工智能得到了实质性或技术改进，或者人工智能被用于解决技术问题，那么就可能被认为具有可专利性。³⁰

97. 在巴西，国家工业产权局（INPI）于2020年更新了计算机实施发明（CII）指南，以处理与人工智能相关的专利申请。³¹专利适格性的确定取决于要求保护的主体是否落入 INPI 审查指南中的可专利性排除范围。修订后的指南还强调，人工智能技术，包括机器学习和深度学习工具，在应用于解决技术问题时，可被视为发明。³²在此，如果人工智能模型或算法脱离了某一特定技术领域的申请，则被视为数学方法或算法，因此不能被授予专利。相比之下，可授权的主体包括与人工智能核心技术变化有关的发明，如改变训练过程或开发新的神经网络架构，前提是如果考虑到所要解决的具体技术问题，这种变化是合理的，此外还要被纳入某一技术领域并产生技术效果。在巴西，神经网络架构单独来看被视为数学方法。关于基于硬件的人工智能系统，为确保专利适格性，必须说明硬件实施的具体特征和细节，仅提及可能的硬件解释是不够的。³³

98. 在中国，修订后的专利审查指南于2024年生效，修改了与人工智能发明相关的审查标准。³⁴一般来说，根据新的标准，如果算法与计算机系统的内部结构具有特定的技术关系，并且能够解决如何提高硬件计算效率或改进执行结果的技术问题，那么人工智能相关发明可以作为可授权主体。³⁵要构成可授权主体，权利要求除了包含规则和商业方法的算法特征外，还必须包含技术特征。更新后的审查指南还包括对大数据处理、深度神经网络等多个人工智能相关领域的主题进行评估的示例。

99. 到目前为止，德国还没有确立专门针对人工智能相关发明可专利性的判例法。不过，由于这些发明在概念上类似于计算机实施发明，因此，通常采用德国联邦法院（BGH）根据德国专利法第1、3和4条制定的程序相关发明三步审查法来处理专利适格性问题。³⁶

²⁹ 例如，见智利、立陶宛和葡萄牙答复 C.9199 的评论意见。

³⁰ 见澳大利亚答复 C.9199 的评论意见。

³¹ 2020 年决议 INPI/PR 第 411 号。

³² 见巴西答复 C.9199 的评论意见。

³³ 同上。

³⁴ 专利审查指南（2023 年）修改解读（四）——涉及计算机程序的发明专利申请审查：
https://www.cnipa.gov.cn/art/2024/1/18/art_2199_189877.html。

³⁵ 同上。

³⁶ 见德国答复 C.9199 的评论意见。

100. 日本特许厅发布的专利和实用新型审查指南中也包含与人工智能发明有关的示例。³⁷这些示例说明了判定人工智能发明是否具有技术性的标准，而技术性是获得专利适格性的关键。日本特许厅强调，人工智能相关发明必须产生特定技术效果或有助于解决技术问题，才能获得专利授权，而没有这种技术贡献的单纯算法或抽象概念则不能授予专利。

101. 在大韩民国，韩国知识产权局（KIPO）编拟了一份详细的人工智能领域审查指南。³⁸与其他局类似，该指南指出，人工智能相关发明的专利适格性标准原则上与计算机软件相关发明的标准相同。一般来说，人工智能发明在以下情况下符合授予专利的条件：(i) 信息处理是通过“软件和硬件的组合”进行实施；(ii) 权利要求不涉及“人类智力活动或离线活动”。韩国知识产权局编拟的审查指南还包括与人工智能发明有关的示例。

102. 新加坡同样为人工智能相关发明的专利适格性判断提供了具体指导。³⁹根据审查指南，人工智能和机器学习方法，如神经网络、支持向量机、判别分析、决策树和 k-means，被视为数学方法，其本身不被视为发明。⁴⁰仅仅使用传统计算机硬件实施人工智能方法不太可能符合适格性标准，除非实际贡献超出了基础数学方法的范围。不过，应用人工智能解决具体问题的人工智能相关发明，如识别人类语言或图像，可能被视为符合专利适格性。通过在问题与数学方法的步骤之间建立充分的联系，权利要求应在功能上明确或隐含地限制在解决具体问题上。⁴¹例如，具体说明一连串数学步骤的输入和输出如何与问题相关，可确保该方法与解决问题之间存在因果关系。此外，提及计算机实施或使用计算机硬件的要求保护的人工智能方法必须证明它们解决了一个具体问题。如果申请仅仅涉及使用传统硬件来实施机器学习方法，那么除非硬件交互实质性地解决了一个具体问题，否则就不太可能符合专利适格性。

103. 在联合王国，上诉法院最近处理了人工神经网络（ANN）的主题适格性问题。⁴²本案中的发明通过向经过训练的 ANN 传递音乐曲目向用户生成音乐曲目推荐。在此，法院认定，计算机被定义为“处理信息的机器”，计算机程序由“向计算机发送的一组行事指令”组成，特别是以特定方式处理信息的指令。⁴³法院进一步得出结论，无论 ANN 是通过硬件还是软件实施，都属于计算机，其权重和偏差也被视为计算机程序。⁴⁴法院裁定，ANN 所做推荐的改进并不构成技术效果，因为“使推荐文件值得推荐的是其语义质量”，而语义质量具有主观性和认知性，不会使该系统成为一个在排除主题之外产生技术效果的系统。⁴⁵因此，本案中的 ANN 发明作为计算机程序本身不具有可专利性。针对这一裁决，英国知识产权局（UKIPO）表示，为主题适格性的目的，今后将像处理任何其他计算机实施发明一样处理 ANN 实施发明。⁴⁶

³⁷ 专利和实用新型审查手册附件 A。在专利适格性方面，讨论的示例包括：指向仅仅是信息呈现的数据的权利要求；能够进行信息处理的数据结构，这种处理可以在语音交互系统中进行；用于分析住宿场所声誉的经训练的模型。

³⁸ 韩国知识产权局人工智能领域审查指南：
<https://www.kipo.go.kr/upload/en/download/Examination%20Guide.pdf>。

³⁹ IPOS 专利申请审查指南，第 8.22-8.27 段。

⁴⁰ 同上。

⁴¹ 同上。

⁴² 专利、外观设计和商标局长诉 Emotional Perception AI Limited [2024] EWCA Civ 825。

⁴³ 同上，第 61 段。

⁴⁴ 同上，第 68 段。

⁴⁵ 同上，第 79 段。

⁴⁶ 《涉及人工神经网络的专利申请法定审查指南》：<https://www.gov.uk/government/publications/examining-patent-applications-involving-artificial-neural-networks/examining-patent-applications-involving-artificial-neural-networks>。

104. 在美利坚合众国，美国专利商标局（美国专商局）发布的 2019 年专利主题适格性指南修订版包含一个示例，具体讨论了一种训练神经网络进行面部识别的计算机实施方法的专利适格性，该方法包括一系列训练步骤。⁴⁷随后，美国专商局于 2024 年更新了人工智能发明专利适格性指南。更新后指南的核心是适用最高法院对于“Mayo、Myriad 和 Alice”案判决的框架。

105. 美国专商局主题适格性分析的步骤 1 处理的是要求保护的发明是否是一种新的和有用的方法、机器、产品、物质组成或其改进（即法定主题）。如果一项发明不属于方法、机器、产品或物质组成，则该发明在法律上不可授予专利。步骤 2A 第一项涉及判断一项权利要求是否指向可专利性的法定例外，如抽象概念、自然法则或自然现象。如果权利要求没有记载法定例外，则认为符合专利适格性，专利适格性分析就此结束。如果确定存在法定例外，步骤 2A 第二项就会评估权利要求项目单独或组合在一起是否明显超出法定例外，从而使其转变成为符合专利适格性的申请。如果权利要求“指向”法定例外，则分析进入步骤 2B，评估要求保护的附加项目是否明显超出所记载的法定例外本身。步骤 2B 包括考虑附加项目是否属于充分了解、常规、传统的活动。基于一个或多个法定考虑因素（例如，认为附加限制是无关紧要的额外解决活动或为适用例外而提供的更多指示），一项权利要求可能会被认定为未明显超出（因此不合格）。相比之下，如果美国专商局的工作人员在步骤 2B 中判断附加项目确实明显超出法定例外，则权利要求符合专利适格性。

106. 关于人工智能发明，更新后的指南特别对上述测试的两个方面作了进一步澄清：(1) 在步骤 2A 第一项中对权利要求是否记载了抽象概念的评估；(2) 在步骤 2A 第二项中对改进考虑因素的评估。更新后的指南提供了大量评注和假设性示例，说明如何将测试中的这些步骤适用于人工智能相关技术。⁴⁸虽然新指南强调分析主题适格性的“Alice/Mayo 测试”并未发生变化，但新指南引入了三个新的或是通过或是未通过测试的示例。这些示例涉及“使用神经网络进行异常检测的人工智能”、“语音信号分析人工智能”和“个性化医疗人工智能”。⁴⁹

107. 在 2018 年 11 月版的欧洲专利局（欧专局）审查指南中，在关于数学方法和方案、从事智力活动、游戏或从事商业活动的规则和方法的章节中，新增了尤其是关于人工智能和机器学习的小节，以更为清晰地确定相关的专利性标准。⁵⁰

108. 关于人工智能技术协助创造的发明，关于可授权主题的考量明显取决于最终发明的性质及其权利要求的撰写方式。例如，在植物被排除在可授权主题之外的国家，对在人工智能工具协助下创造出来的创新型新植物进行了界定的专利权利要求不具有可专利性。

⁴⁷ 2019 年修订的专利主题适格性指南，例 39。

⁴⁸ 2024 年关于专利主题适格性（包括人工智能）的指南更新：
<https://www.federalregister.gov/documents/2024/07/17/2024-15377/2024-guidance-update-on-patent-subject-matter-eligibility-including-on-artificial-intelligence>。

⁴⁹ 2024 年 7 月主题适格性示例：<https://www.uspto.gov/sites/default/files/documents/2024-AI-SMEUpdateExamples47-49.pdf>。

⁵⁰ 欧洲专利局（欧专局）审查指南 G 部分第二章 3.3.1。该指南的总体观点是，人工智能和机器学习基于计算模型和算法来分类、聚类、回归和降维，这些操作本身具有抽象的数学特性，无论它们能否基于训练数据“进行训练”。但是，如果人工智能和机器学习被应用于各种不同的技术领域，作出技术贡献并支持技术目的的实现，则这样的发明可视为可授权主题。

C. 新颖性和创造性

109. 据说创造性分析是专利性标准中最难评价的要求。⁵¹在被驳回的专利申请中，很多申请的驳回理由是缺乏创造性。当专利的有效性遭到第三方质疑时，它们经常把与创造性要求不符作为其主张的依据。人工智能领域的专利申请和专利似乎面临着同样的问题。已出现了一些问题，即对于涉及使用人工智能的发明，应如何分析创造性步骤以及界定本领域技术人员这一概念。⁵²尽管可用的数据有限，但在第三方提出的有关人工智能相关申请/专利的异议中，很多是以缺乏创造性（显而易见性）为理由提出的。⁵³

110. 当新技术出现时，创造性评价经常面临特别的挑战。这是因为现有技术参考文献很少，假定的本领域技术人员的确切范围以及该领域的公知常识还未充分确定。缺少判例法和正式指南使主管局难以对创造性作出一致的评价。但是，随着技术趋于成熟，在很多技术领域已逐渐出现了普遍解释和标准做法。

111. 由于创造性评价是由本领域技术人员作出，因此确定这一假定人员所掌握的知识技能水平就成为创造性评价的基石之一。⁵⁴在每个具体个案中，这一知识技能的确切水平都要得到界定。它也随着技术发展而发生变化。一般来说，假定的本领域技术人员的能力和知识可酌情对应从事各个相关领域技术人员团队的水平。⁵⁵因此，预期的情况是一个人工智能工具在相关领域使用得越多，这种使用的创新性就越低，因为本领域技术人员，即能够使用该人工智能工具的跨学科团队，会在其研究中转而使用这种工具。类似的考量也适用于“公知常识”的概念。⁵⁶

112. 在中国，对涉及算法特征或商业规则和方法的发明专利申请进行新颖性审查时，应当考虑权利要求中说明的所有特征。这些特征既包括技术特征，也包括算法特征或商业规则和方法特征。⁵⁷此外，如果与计算机系统的内部结构有特定技术关系的算法可提高系统的内部性能，那么在评估创造性时也应考虑该算法。同样，如果一个解决方案可通过技术特征，或通过技术特征、算法特征和商业规则或方法的交互作用改善用户体验，那么在评估创造性时也应考虑这种用户体验的改善。⁵⁸例如，如果权利要求中的算法应用于特定的技术领域，能够解决特定的技术问题，则可以认为算法特征与技术特征在功能上相互支持，具有交互关系，并且算法特征成为所采用的技术手段的一部分。⁵⁹

113. 日本特许厅发布的专利和实用新型审查手册附件 A 载有一些关于评价人工智能相关发明创造性的示例。⁶⁰例如：

- 缺乏创造性，因为发明只是在人工智能系统中把人的操作进行了系统化处理（示例 33）；
- 缺乏创造性，因为发明只是对一种根据输入数据预测输出数据的方法进行了修改（示例 34）；

⁵¹ 更多关于创造性要求如何在不同国家实施的信息见 SCP/22/3、SCP/28//4、SCP/29/4 和 SCP/30/4。

⁵² 见智利答复 C. 9199 的评论意见。

⁵³ 2019 年世界知识产权组织技术趋势——人工智能，第 115 页至第 117 页。

⁵⁴ 见文件 SCP/22/3。

⁵⁵ 文件 SCP/22/3 第 34 段和第 35 段。

⁵⁶ 见文件 SCP/28/4。

⁵⁷ 中国国家知识产权局审查指南第二部分第 9 章第 6.1.3 条。

⁵⁸ 专利审查指南（2023 年）修改解读（四）——涉及计算机程序的发明专利申请审查。

⁵⁹ 日本特许厅-中国国家知识产权局人工智能相关发明比较研究。

⁶⁰ 专利和实用新型审查手册附件 A，示例 31 至 40，日本特许厅。

- 由于添加某些训练数据会产生显著效果而涉及的创造性（示例 34）；
- 缺乏创造性，因为修改机器学习的训练数据只是已知数据的组合，并未产生任何显著影响（示例 35）；
- 由于对训练数据进行某种预处理而涉及的创造性（示例 36）；
- 缺乏创造性，因为发明仅仅是利用生成式人工智能对人类任务进行系统化处理（例 37）；
- 根据生成式人工智能应用的特征而涉及的创造性（例 38）；
- 根据从输入数据估算输出数据的经过训练模型的学习方法差异而涉及的创造性（例 39）；及
- 根据在利用人工智能将人类任务简单系统化的基础上增加新特征而涉及的创造性（例 40）。

114. 欧洲专利局在创造性方面对人工智能相关发明的评估方式与其他计算机实施发明的评估方式类似。在此，要确保在评估创造性时只考虑有助于实现发明技术特性的技术特征。特别是“非技术性”特征，在这里指的是那些本身属于被排除在可专利性之外的特征，只有当这些特征有助于解决技术问题时，才考虑对其进行评估。如果技术特征对于解决技术问题不发挥作用，那么在评估创造性时也可能不予考虑。欧专局上诉委员会曾多次做出裁决，涉及与人工智能和机器学习相关发明的创造性评估。⁶¹

115. 关于人工智能机器“创造”的发明，存在关于人工智能机器大量创造“新发明”的关切，有人担心这会造成所有发明都是由机器创造并获得专利授权的情况。在一定程度上反映了上述关切的一些项目通过公开人工智能机器的输出来产生使用人工智能技术的“现有技术”，使他人无法为上述任何输出获得专利权。⁶²

116. 关于新发明，充分公开要求和工业实用性（实用性）要求可避免对这样的发明授予专利，例如只是已知化学元素的组合，并未说明如何生成这种合成物及其如何使用。同样地，在已公布的引文中所述的信息只能被视为公众可用的信息，因此是适格的现有技术引文，条件是对这些信息进行了充分细致的说明，使本领域技术人员能够实现这些教导。例如仅仅以化学式形式公开的化学结构很有可能不被视为适格的否定对应化学合成物新颖性/创造性的现有技术引文。

117. 创造性（非显而易见性）要求的出发点是应对可作为公知常识的明显结果推断出的发明给予专利保护，因为它对于社会的贡献微乎其微。⁶³这样的政策目标可指导在包括人工智能相关发明在内的个案中对创造性进行判定。

D. 充分公开和权利要求⁶⁴

118. 与创造性评价相同，新技术对于发明的公开提出了特别的挑战，要求清楚完整地公开发明，以及清晰扼要地撰写权利要求，以充分涵盖法律保护的范围。缺少判例法和正式指南使知识产权局和专利制度的用户同样难以对是否符合公开要求进行评价。

⁶¹ 《人工智能和机器学习判例法》：https://www.epo.org/en/legal/case-law/2022/clr_i_d_9_2_11_e.html。

⁶² 所有现有技术项目（<https://allpriorart.com/about/>）。

⁶³ 文件 SCP/22/3 第 3 段。

⁶⁴ 另见文件 SCP/34/5（关于充分公开的进一步研究（第一部分）），四、人工智能相关发明。

119. 关于要求保护发明的说明书，一般来说，国家/地区专利法要求专利申请人应以充分清晰完整的方式对发明进行公开，使本领域技术人员能够实现要求保护的发明（充分公开要求）。⁶⁵专利制度正是通过这一要求为传播信息以及获取专利申请和专利所包含的技术知识提供便利。这使技术知识在公众中的积累得以扩大，并增进了社会总体效益，例如促进技术转让和避免重复性的研发工作。

120. 在人工智能技术方面，一个可能的问题是人工智能算法、训练模型、神经网络架构、学习过程、训练数据、硬件组件等信息在专利申请中应以怎样的程度公开才能符合充分公开条件。⁶⁶因此，评估人工智能发明是否充分公开是一个新的挑战，许多国家尚未就此确立判例法。⁶⁷挑战之一可能是由于就目前的深度学习技术而言，人们难以识别深度学习神经网络中所执行的每个流程步骤，以及明确解释神经网络如何得到最终结果。如果一个系统有数千万个用于分类的权重，则该系统会过于复杂，以至于无法以人可以理解的方式对其进行表述。在某些情况中，在没有实际试验数据的情况下对人工智能输出进行合理解释（即以可信的方式进行论证）可能更为困难。

121. 与此同时，要求保护的发明在专利申请说明书部分的公开程度显然取决于在申请的权利要求部分中要求保护的内容。例如一项发明是关于通过利用特定数据集训练深度学习算法来应用人工智能技术解决一个问题，如果要求保护的发明有更广泛的应用，则可能要求在说明书中公开本领域技术人员实现要求保护的发明广泛应用所必需的所有数据集类型，而不是只公开一种数据集类型。

122. 在这方面，本领域技术人员的概念对于评价充分公开同样重要。例如，如果一项人工智能技术被应用于某一领域的发明（如图像识别神经网络应用于安防和监控领域的发明），则人工智能技术领域和监控领域技术人员团队可作为假定的本领域技术人员来对这项发明进行评价。

123. 另一个问题可能是由于深度学习技术的非确定性：它们需要进行一些随机的初始化。因此，相同的训练数据和相同的神经网络架构也可能使机器学习有略微不同的绩效。使用相同训练数据和相同神经网络架构对模型进行训练会产生两种略有区别的训练行为。与生物材料不可避免地存在生物差异性的情况相同，可以基于专利申请中的公开考虑要求保护的发明所谓的可再现性和可行性。

124. 关于训练数据，通过一项特定人工智能技术解决一个问题可能需要一个特定数据集。训练数据集在深度机器学习的绩效方面所发挥的重要作用可能会让人们提出这样的问题，即此类数据集在专利申请中的公开程度应为多少，以及是否应提供此类数据集以供第三方对要求保护的发明进行验证（即要求保护的发明是否确实起作用）。

125. 关于权利要求，很多国内法规定权利要求应清晰扼要。此外，权利要求应得到说明书的支持（支持要求）。⁶⁸一般来说，这项要求的出发点是要求保护的发明不应超出说明书中所公开披露发明的范围。同样地，美国法律所规定的书面说明书要求的关键政策目标⁶⁹是“明确传达申请

⁶⁵ 见文件 SCP/22/4。另见“国家/地区专利法的若干方面——充分公开”：
https://www.wipo.int/scp/en/annex_ii.html。

⁶⁶ 见智利和德国答复 C.9199 的评论意见。

⁶⁷ 例如，见德国答复 C.9199 的评论意见。

⁶⁸ 见文件 SCP/22/4。

⁶⁹ 美国法案第 35 章 112(a) 条。见文件 SCP/22/4。

人发明了要求保护的主题的信息，以及使公众掌握申请人要求保护发明的信息”。⁷⁰因此，这些要求体现了这样的基本原则，即对于申请人截至申请日并未发明的内容以及截至申请日并未通过专利申请中的公开与公众分享的内容不能给予专利保护。由于人工智能相关发明大多是计算机实施发明，在人工智能相关发明的权利要求撰写方法方面，申请人可能面临相同的挑战，即在权利要求中要为其发明确定适当的保护范围。

126. 关于人工智能相关发明对于公开要求的适用，日本特许厅发布的专利和实用新型审查手册附件 A 载有若干具体示例。⁷¹这些示例主要阐释了这样的情况，即人工智能技术被应用于各种不同技术领域的发明，因此机器学习通常要求提供多种类型的训练数据。它们讨论了为了符合公开要求而显示在这些数据中间存在某种关系（如相关性）的重要性。此外，有一个示例讨论了假定人工智能技术为要求保护的产品发明提供某种功能的情况。要求保护的发明没有满足公开要求，因为说明书中只提供了人工智能推理数据（没有产品的试验数据），而现有技术和公知常识都没有提出人工智能推理数据可以取代试验数据。

127. 在大韩民国，说明书应详尽解释人工智能相关发明的具体手段、解决的技术问题和提供的解决方案，确保技术人员能够轻松理解和再现该发明。如果一项发明没有具体规定经过训练的模型的输入数据和输出数据之间的相关性作为实现人工智能相关发明的特定手段，则该发明被视为公开不充分。经过训练的模型的输入数据和输出数据之间的相关性是指：(i) 具体说明训练数据；(ii) 要求保护的发明各特征之间存在解决技术问题相关性；(iii) 具体说明一种通过使用训练数据进行训练的学习模型或一种训练方法；(iv) 通过这种训练数据和训练方法生成用于解决要求保护的发明的技术问题的训练模型。然而，如果本领域技术人员，在将申请提交时的公知常识考虑在内的情况下，能够根据发明说明书中所述的模型推测或理解这种相关性，则视为满足了能够实现的要求。此外，仅使用功能框图或一般流程图而不提供具体实施细节通常是不够的。申请必须说明硬件或软件如何以本领域技术人员可以理解和再现的方式实现发明的功能。这种详细的公开可确保发明能被有效地再现和利用，从而满足能够实现要求，并支持创新的实际应用性。⁷²

128. 2024 年 3 月，欧洲专利局更新了审查指南，对人工智能相关发明的充分公开要求作出了实质性修改。⁷³更新后的指南中的一项重要修订是，要满足充分公开要求，可能必须在专利申请中提供关于要求保护的人工智能模型的训练数据的详细信息。其目的是确保充分公开，并作为技术效果的证据。具体来说，更新后的指南指出，下列情况下专利公开不充分，即如果“数学方法和训练数据集的公开不够详尽，无法在要求保护的整个范围内再现技术效果。缺乏详细信息的公开可能使得专利公开更像是研究计划进行征集”。⁷⁴新指南还指出，“机器学习算法实现的技术效果可能是显而易见的，或是通过解释、数学证明、实验数据或类似方法所确立[……]如果技术效果取决于所使用的训练数据集的特定特征，则必须公开重现技术效果所需的这些特征，除非技术人员能够利用公知常识确定这些特征，而不会造成不必要的负担。不过，一般来说，没有必要公开具体的训练数据集本身”。⁷⁵尽管该指南目前规定，一般不要求完全公开基础训练数据集，但在哪些情况下必须公开数据集尚未完全确定。

⁷⁰ 同上。

⁷¹ 专利和实用新型审查手册附件 A，示例 46 至 51，日本特许厅。

⁷² 《人工智能领域审查指南》，韩国知识产权局。

⁷³ 《欧洲专利局（欧专局）审查指南》，G 部分，第二章，3.3.1。

⁷⁴ 《欧洲专利局（欧专局）审查指南》，F 部分，第三章，公开不充分。

⁷⁵ 《欧洲专利局（欧专局）审查指南》，G 部分，第二章，3.3.1。

129. 欧洲专利局做出了两项与充分公开和人工智能有关的决定。在 T 0161/18 案中，欧洲专利局上诉委员会认定，该案中涉及人工神经网络的专利申请不符合充分公开的要求。上诉委员会指出，该申请没有公开哪些输入数据适合训练人工神经网络，也没有公开至少一个适合解决该技术问题的数据集。⁷⁶同样，在 T 1191/19 案中，一件专利申请也被认定不符合充分公开的要求。在得出这一结论时，委员会注意到该申请没有提供任何训练数据的示例，甚至没有提供使该发明能够提供有意义的预测所需的训练数据量。⁷⁷这些案例强调，申请人必须确保其公开的信息能够实现，可以包括详细的训练数据和其他相关信息，以符合欧洲专利局的要求。

130. 作为一个新的技术领域，对人工智能相关发明适用充分公开要求仍在演变之中，对此尚未有明确的规定。多个司法管辖区正在制定报告和规则，以应对这些特殊的挑战。⁷⁸该领域立法、指南和法院裁决的不断发展将有助于建立更为精准的标准和实践。从一般意义上讲，许多人工智能算法固有的黑匣子现象可能不会对发明的充分公开造成问题，只要提供足够的详细信息来说明使用哪种人工智能算法以及如何对其进行训练即可。⁷⁹

E. 工业实用性

131. 关于要求保护发明的可再现性和可行性，在一些国家，符合工业实用性要求可能也意味着要求保护的发明在必要时必须能够再现且再现结果具有相同的特性。⁸⁰

132. 工业实用性的一个关键方面是要求保护的发明的可再现性和一致性。对于人工智能系统，特别是那些应用深度学习和神经网络的系统来说，由于其固有的复杂性以及决策过程不透明性，要证明绩效的一致性可能具有挑战性。

133. 举例来说，用于工业机械预测性维护的人工智能系统只有在能够持续准确预测故障的情况下，才可被视为具有工业实用性。这可能需要详细记录人工智能模型，包括其训练数据、算法和验证结果，以确保它能在工业环境中可靠实施并产生一致的结果。

F. 发明人权利和所有权

134. 随着人工智能日益融入发明过程，一些重要问题开始显现，涉及人工智能创造之发明的所有权，以及在有人工智能协助的情况下自然人列为发明人的标准。⁸¹SCP 在 2023 年制定了文件 SCP/35/7，以解决在人工智能相关发明的背景下不断演变的发明人权利和所有权问题。该文件深入概述了人类与人工智能在发明过程中的交互关系，并透彻分析了当前围绕人工智能发明人权利的法律框架和政策问题。⁸²因此，本节仅概述关于人工智能相关发明的发明人权利/所有权问题。

专利法规定的发明人权利和所有权

135. 《巴黎公约》第 4 条之三规定，发明人有权要求在专利中提及自己是发明人。该条是指发明人享有在巴黎联盟的所有国家在为其发明授予的专利中记录自己是发明人的权利，通常称之为署名权。普遍的理解是发明人可以放弃这项权利，除非国家立法另有规定。由于《巴黎公约》没有

⁷⁶ 欧洲专利局上诉委员会第 T0161/18 号决定。

⁷⁷ 欧洲专利局上诉委员会第 T1191/19 号决定。

⁷⁸ 见巴西答复 C. 9199 的评论意见。

⁷⁹ 见德国和葡萄牙答复 C. 9199 的评论意见。

⁸⁰ SCP/5 非正式文件（工业实用性/实用性要求在国家和地区法律下的实际应用）。另见欧洲专利局上诉委员会判例法，I. E. 2. 部分。

⁸¹ 见智利答复 C. 9199 的评论意见。

⁸² 委员会在 SCP 第三十五届会议上商定，秘书处将更新文件 SCP/35/7 第五和第六部分并提交 SCP 第三十七届会议。

给出“发明人”一词的定义，因此发明人的确定以及行使该署名权的程序由各成员国在其适用法中作出规定。⁸³如果一名以上的发明人共同创造了一项发明，则他们是共同发明人。

136. 尽管可专利性要求（如可授权主题、新颖性、创造性（显而易见性）、工业实用性（实用性）和公开要求）不涉及发明人权利的问题，但不准确填写发明人姓名可能产生严重的法律后果。发明人指定不准确会导致各种法律后果，不同司法管辖区对这种不准确的问题提供救济措施。

137. 尽管并非所有国内法都规定了“发明人”的定义，但考虑到专利制度的出发点以及署名权是与专利权相关的基本权利之一，可能存在的普遍推定是专利法下的（一个或多个）发明人是指（一个或多个）个人。⁸⁴如果这是一个正当推定，那么逻辑结论可能就是无论人工智能机器为发明的构想作出了多大程度的贡献，机器不能作为发明人。

138. 如果在发明创造的过程中使用了人工智能系统，只要在这一过程中有一人（或多人）符合适用法下“发明人”的条件——宽泛地说就是为构想要求保护的发明作出了贡献——则该人（或这些人）就是该发明的发明人，发明人可以是人工智能程序员、人工智能开发人员、人工智能用户或其他人员。这里理论上存在的一个问题是如果没有人符合适用法下发明人的条件，谁享有对于专利的权利？

139. 尽管随着技术的进步，预计人工智能机器将具有更强的认知能力，但技术发展往往是增量式发展。此外，人工智能技术在发明创造过程中根据不同的情况可能发挥不同的作用，即可能发挥从只是作为辅助工具到帮助构想发明概念的手段的各种作用。因此，把“人类创造的发明”和“机器创造的发明”对立起来对于关于发明人权利问题的复杂讨论来说似乎过于简单。

140. 一般来说，专利权在一开始属于发明人，而发明人可以把权利转让给其他自然人或法人。因此，发明人权利（即发明了发明的人）和所有权（即有权提交专利申请或获得专利的人）是两个相互独立但又相互联系的概念。在很多国家，如果发明属于职务发明，专利权原则上属于员工，对此通常附带某些条件。⁸⁵因此，发明人权利/所有权问题可作为设计专利制度过程中关键政策问题的一部分。

人工智能发明人权利和 DABUS 案

141. Stephen Thaler 提交了两件专利申请，其中发明人姓名填写的是人工智能系统“Device for the Autonomous Bootstrapping of Unified Science（统一科学自主引导装置）”（DABUS）。这两件申请最初是向欧洲专利局（欧专局）和英国知识产权局（UKIPO）提交，据说随后又在其他 15 个司法管辖区提出了申请。产权组织国际局收到了一件根据专利合作条约（PCT）提交的将 DABUS 列为发明人的国际申请（PCT/IB2019/057809）。

142. 多个知识产权局通过进入国家阶段的 PCT 申请或直接提交的方式收到了一件或一件以上的 DABUS 申请。已受理这些申请的知识产权局大多以未在发明人姓名中写明自然人姓名为由驳回了这些申请。在许多情况下，申请人就这些决定向法院提出上诉，法院根据专利法拒绝将人工智能机器作为发明人。

⁸³ 《保护工业产权巴黎公约申请指南》，G. H. C. Bodenhausen（产权组织出版物 No. 611）。

⁸⁴ 根据 35 U.S.C. § 100(f)， “发明人”是指“发明或发现了发明主题的个人或是，如果是共同发明，全部个人”。在美国，要求保护发明的发明人或是共同发明人中的每个人原则上必须对申请进行宣誓或作出声明。

⁸⁵ 为了完整的目的，还应补充的一点是专利权也可以通过继承转让给他人。

143. 文件 SCP/35/7 概述了澳大利亚、巴西、加拿大、德国、印度、新西兰、大韩民国、南非、联合王国、美利坚合众国的知识产权局和法院（如有）以及欧洲专利局的决定。

四、作为工具应用于专利制度的申请过程和管理工作的人工智能技术

144. 人工智能技术解决方案可被用于专利程序以及专利程序以外的情况中，即作为申请人提交专利申请的工具、专利局处理专利申请的工具、专利权人维权的工具、第三方提出宣告专利无效请求的工具或司法机关解决争议的工具等。

A. 知识产权主管部门的工具

145. 知识产权局已开始利用人工智能技术为知识产权管理和交付提供便利。产权组织知识产权局人工智能倡议索引⁸⁶是一个在线门户，在该网站上按国家/领土和人工智能商业应用开列有人工智能技术在各知识产权局的使用情况。索引中的商业应用类别（是人工智能技术为知识产权局提供便利的主要业务领域）有：(i) 数字化和流程自动化；(ii) 专利审查管理；(iii) 问询处服务；(iv) 图像检索；(v) 机器翻译；(vi) 专利分类；(vii) 专利现有技术检索；(viii) 商标分类；(ix) 版权登记；及(x) 数据分析。

146. 在 2018 年 5 月 23 日至 25 日在日内瓦举行的产权组织 ICT 战略和人工智能促进知识产权管理知识产权局会议期间，会议主题之一是知识产权局如何已经利用以及可以利用人工智能和其他先进技术的应用。⁸⁷会议讨论表明多个主管局在把人工智能的潜力用于知识产权管理系统方面取得了进展，并且显示各主管局希望持续交流有关人工智能的信息和经验，这么做还可避免尤其是进行重复性工作。⁸⁸作为会议的一项跟进工作，产权组织为人工智能创建了一个专门网页⁸⁹，并为讨论 ICT 战略和人工智能促进知识产权管理建立了电子论坛，仅对知识产权局所指定的专家开放。此外，产权组织标准委员会（标准委）成立了 ICT 战略和标准工作队，负责特别是对提交会议的建议进行审查。⁹⁰

147. 此外，自 2019 年 2 月以来，产权组织知识产权与人工智能对话（后更名为产权组织知识产权与前沿技术对话）提供了一个开放、包容的论坛，就人工智能等前沿技术对知识产权的影响，与尽可能广泛的利益攸关方进行接触，并促进他们之间的讨论和知识积累。⁹¹最近，2022 年举行的第六届会议重点讨论了人工智能发明以及世界各地的知识产权局如何支持人工智能，2023 年举行的第八届会议讨论了生成式人工智能对内容创作的影响以及与此相关的众多知识产权问题。

148. 在专利管理领域，国家和地区专利局开发了（或正在开发）人工智能应用工具用于：专利申请分类；手续检查、现有技术检索；相关文件的机器翻译；协助实质审查（如自动标注专利文献以及自动检测不属于可授权主题的主题）；以及更为普遍的数据转换和文件管理。⁹²

149. 产权组织国际局也在工作中使用人工智能，以强化本组织的职能和程序。国际局还开发并提供了一系列人工智能驱动的服务和工具，为用户和利益攸关方提供帮助。产权组织目前在多个

⁸⁶ https://www.wipo.int/about-ip/en/artificial_intelligence/search.jsp。

⁸⁷ 会议的文件和演示文稿见：https://www.wipo.int/meetings/en/details.jsp?meeting_id=46586。

⁸⁸ 文件 WIPO/IP/ITAI/GE/18/5（协调员的总结）。

⁸⁹ https://www.wipo.int/about-ip/en/artificial_intelligence/。

⁹⁰ 文件 CWS/6/3。

⁹¹ https://www.wipo.int/about-ip/en/frontier_technologies/frontier_conversation.html。

⁹² 产权组织知识产权局 AI 倡议索引。另见文件 SCP/32/4 和 4 Corr. 以及 SCP/34/4。

领域使用人工智能，包括语音转文字、全球品牌数据库内的图像检索、自动专利分类、尼斯和维也纳分类协助以及机器翻译（WIPO Translate）。⁹³

150. 在 SCP 第三十一届、第三十三届和第三十五届会议的交流会上，各知识产权局和产权组织国际局介绍了它们在专利管理和专利审查中使用人工智能的情况。⁹⁴

B. 申请人、第三方和知识产权专业人员的工具

151. 考虑到通过专利制度产生了越来越多向公众公开的信息，人工智能技术还可以帮助申请人、第三方和知识产权专业人员提升各自活动的质量和效率。

152. 国际保护知识产权协会（AIPPI）、美国知识产权法律协会（AIPLA）和国际知识产权律师联合会（FICPI）认为，人工智能在知识产权实务中的应用可分为三类：(i) 文件自动化；(ii) 流程自动化；及(iii) 人工智能产生的深入分析。⁹⁵它们预计基于人工智能的文件自动化能够结合语境分析表述，并帮助例如申请撰写和校对。基于人工智能的流程自动化可利用专利数据进行检索，并被用于建立任务列表、生成审查意见通知书壳以及创建和管理信息披露声明。人工智能产生的深入分析向专利制度的用户提供分析和预测，用户可以通过这些分析和预测更好地作出知情决定。

[后接附件]

⁹³ 欲了解详细信息，请访问产权组织网站：<https://www.wipo.int/en/web/ai-tools-services/>。

⁹⁴ 演示文稿可在各 SCP 会议网页上查阅。文件 SCP/32/4 和 4 Corr. 以及 SCP/34/4 分别载有 SCP/31 和 SCP/33 期间举行的交流会的报告。SCP 第三十六届会议期间也将举行类似的交流会。

⁹⁵ AIPLA/AIPPI/FICPI 人工智能入门讨论会：AIPPI/AIPLA/FICPI 人工智能交流讨论会，2019 年 3 月 28 日和 29 日，<https://ficpi.org/colloquium>。

产权组织和各知识产权局发布的有关人工智能的网页和出版物

人工智能与发明人资格（文件 SCP/35/7）

https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=620584

关于公开充分性的进一步研究（第一部分）（文件 SCP/34/5）

https://www.wipo.int/meetings/en/doc_details.jsp?doc_id=582853

“知识产权与前沿技术”网页

https://www.wipo.int/about-ip/en/frontier_technologies/

- 创新生态系统为人工智能做好准备——知识产权政策工具包

2019 年产权组织技术趋势——人工智能

https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf

生成式人工智能，专利态势报告（2024 年）

https://www.wipo.int/web-publications/patent-landscape-report-generative-artificial-intelligence-genai/assets/62504/Generative%20AI%20-%20PLR%20EN_WEB2.pdf

澳大利亚

“生成式人工智能与知识产权制度”网页

<https://www.ipaustralia.gov.au/temp/Generative-AI-and-the-IP-System.html>

巴西

金砖国家知识产权局最终报告，IP BRICS-INPI

<http://www.ipbrics.net/secondpage/project/Patent%20Processes%20and%20Procedure%20-%20AI%20Study%20Report.pdf>

中国

涉及计算机程序的发明专利申请审查

https://www.cnipa.gov.cn/art/2024/1/18/art_2199_189877.html

日本

“人工智能相关发明”网页

<https://www.jpo.go.jp/e/system/patent/gaiyo/ai/index.html>

日本特许厅-中国国家知识产权局人工智能相关发明比较研究（2023 年）

https://www.jpo.go.jp/e/news/kokusai/cn/document/ai_report_2023_e/cn_ai_report_en.pdf

日本特许厅与欧专局计算机实施发明/软件相关发明比较研究（2021 年）

https://www.jpo.go.jp/e/system/laws/rule/guideline/patent/document/ai_jirei_e/01_en.pdf

大韩民国

人工智能领域审查指南 (KIPO)

<https://www.kipo.go.kr/upload/en/download/Examination%20Guide.pdf>

联合王国

人工智能相关专利申请审查指南

<https://www.gov.uk/government/publications/examining-patent-applications-relating-to-artificial-intelligence-ai-inventions/guidelines-for-examining-patent-applications-relating-to-artificial-intelligence-ai>

美利坚合众国

“人工智能” 网页

<https://www.uspto.gov/initiatives/artificial-intelligence>

欧洲专利局

“人工智能” 网页

<https://www.epo.org/en/news-events/in-focus/ict/artificial-intelligence>

[附件和文件完]