数据驱动与隐私防御：探索人工智能公司的信息收集与使用行为

彭泳谦 2100017731

1.摘要

本研究旨在探讨在快速发展的人工智能时代，数据的收集、使用与个人隐私保护之间的关系。随着技术的进步，企业和政府机构越来越依赖大数据分析来提升服务质量和决策效率，但这同时也引发了对个人隐私、知识产权侵犯的担忧。本文将以政府、企业、个人的三重视角，通过文献综述和案例分析，评估当前的数据保护法律框架、技术隐私保护措施的有效性，以及公司的数据使用行为。研究发现，尽管存在多种隐私保护措施，但在实践中仍存在不足，个人隐私、知识产权仍会受到侵犯。本文提出了一系列针对政策制定者、企业和个人的建议，以改善当前的隐私、知识产权保护状况，并为未来的研究提供了方向。

2.引言

在 21 世纪的信息化社会，数据已成为推动经济和技术发展的核心资产。特别是随着人工智能技术的突破和应用，数据的价值得到了前所未有的放大。企业和政府机构通过收集和分析个人数据，能够提供更加个性化的服务，优化产品设计，提升运营效率，并作出更加精准的决策。然而，数据收集活动也引发了广泛的隐私、产权保护问题，如何在促进技术发展和保护个人隐私、知识产权之间找到平衡点，已成为全球性的挑战。

当前，虽然各国都在努力完善相关的数据保护法律法规，例如欧洲的通用数据保护条例(GDPR)[1]和美国加州的消费者隐私法案(CCPA)[2]，但这些法律法规在实际执行中仍面临许多挑战。此外，技术对策如数据加密和匿名化虽然在保护数据安全方面发挥了作用，但仍无法完全解决隐私泄露的问题。其次，对于企业来说，获取大量的数据来进行人工智能模型的训练是一个必须进行而又十分困难的步骤，使用“爬虫”获取的受知识产权保护的数据和用户在使用产品时产生的信息，已经成为行业中的“潜规则”。同时，个人对于某些平台使用存在刚需，这也使他们不得不承担在使用过程中存在的信息泄露风险。

3.人工智能发展概况

3.1 逻辑推理与符号处理阶段

从 1950s 到 1970s，人工智能处于逻辑推理与符号处理阶段，包括之后的二十年，这一段时间内，研究者受存储、运算能力的不足的限制，只能通过基于规则的专家系统，来模拟智能。这段时期的智能系统基本不涉及到数据的使用，也就不会产生隐私侵犯等问题。

3.2 基于统计的机器学习阶段

从 1990s 到 2000s，随着计算能力的增强和互联网的兴起，可用于机器学习的数据量大幅增加。人们开始更多地关注基于统计的学习方法，如支持向量机[3]和决策树[4]等。但此时的互联网还不是人们生活中的必需品，数据挖掘的方法也极其有限，隐私泄漏问题仍未广受关注。

3.3 数据驱动的深度学习阶段

从 21 世纪开始，以 AlexNet[5]在 ImageNet[6]挑战中取得的突破为起点，人工智能进入了深度学习和大数据的时代。深度学习与大数据，成为了人工智能领域的两个支柱。深度学习，通过构建深层的神经网络，能够对大量的数据进行处理，挖掘数据之间的关系，形成高层次的表征。而此时互联网已经逐渐渗透了人们的生活，积累了大量的数据。在二者共同的作用之下，人工智能高速发展，大量的研究者、投资者涌入了人工智能领域，基于人工智能的应用也步入了人们的生活。然而，数据的大量收集和使用，引发了政府对规范数据使用的需要、企业之间围绕数据的激烈竞争、个人对于隐私信息泄漏的担忧。在这种情况下，政府如何制定相关的法律法规，以规范人工智能行业中的数据使用行为；企业如何合理地使用数据；个人如何保护自己的数据，成为了值得讨论的话题。

4.法律框架和政策

4.1 国际隐私保护法律标准

随着全球化的加速，国际有关数据、隐私保护的法律标准逐渐成为各国制定隐私保护法律的重要参考。其中，欧盟的《通用数据保护条例》(GDPR)[1]被认为是目前最严格、最完善的隐私保护法律之一。该条例规定了数据处理的基本原则，包括合法、透明、目的限制、数据最小化、存储限制、完整性和保密性等。此外，该条例还规定了数据处理者的义务和责任，包括告知用户数据处理的目的、收集用户同意的方式、保护用户数据的完整性和保密性等。

4.2 法律对人工智能公司的约束作用

法律对人工智能公司的约束作用主要体现在以下几个方面：

1）合法性约束：人工智能公司的数据处理活动必须符合相关法律法规的规定，否则将面临法律责任和处罚。

2）透明度约束：人工智能公司必须向用户明确告知数据处理的目的、方式和范围，以及数据的使用和存储情况，以便用户能够了解自己的个人信息被如何处理。

3）保密性约束：人工智能公司必须采取必要的技术和管理措施来保护用户的个人信息，防止数据泄露和滥用。

4）责任和义务约束：人工智能公司必须承担起相应的责任和义务，如告知用户数据的处理情况、为用户提供投诉和申诉渠道等。

4.3 法律框架中存在的漏洞

尽管国际和各国在隐私保护法律方面取得了一定的进展，但仍存在一些漏洞和不足。例如，一些法律法规对于人工智能公司的数据处理活动缺乏明确的规定和监管措施；一些国家对于数据的保护标准不一致，导致跨国数据处理存在一定的法律风险；一些法律法规对于用户投诉和申诉的处理机制不够完善等。导致人工智能公司的数据使用行为无法受到有效监管，带来了隐私、知识产权侵犯风险。

5.人工智能模型中的数据

5.1 人工智能模型中的数据

人工智能公司的数据收集是其运营的核心组成部分。人工智能模型的训练和优化需要大量的数据。对于人工智能模型来说，模型参数量的提升，就意味着最终训练完成时，模型性能的提升，尤其是对于自然语言模型来说。而模型参数量提升后，需要同等数量的训练数据量的增加，才能让模型达到最佳的性能[7]。在这样的趋势下，数据成为了每一位人工智能研究者的必需品，成为了人工智能企业的财富和商业机密。

由于目前深度神经网络可解释性的缺乏，端到端的训练使人工智能模型像一个黑匣子，难以对模型本身进行解耦。人工智能模型的这种特点，使回推已训练完成的模型使用到的训练数据，成为了一个几乎不可能完成的任务。在这种情况下，政府部门和监管机构缺乏有效的方法，评估企业和个人在模型训练中是否使用了侵犯著作权或隐私的不合法内容。另外，可解释性的缺乏，还会带来数据收集、使用量和范围不明确的问题，何种数据为什么以及如何带来性能上的提升，尚没有一个解释的方法。这就给为了隐私保护而存在的“最小收集”原则的实施，带来了实践中的困难。

人工智能模型的上述两个特点，很自然地会引出两个矛盾点：模型性能的提升给用户带来的便利、给公司带来的效益，以及使用模型训练数据对于其他个体权益的侵害。另外，数据的形势都是电子存储，由于这种形式的数据的可复制性，一旦发生泄漏，造成的影响巨大，且无法完全回收清除数据。

5.2 数据收集

人工智能公司收集数据的方法多种多样，包括直接从用户处获取、通过第三方购买、利用在线跟踪技术等。例如，社交媒体平台通过分析用户的点赞、评论和分享行为来了解用户的兴趣和网络行为模式。但在这个过程中，用户由于司法意识的缺失，很有可能不会仔细地阅读相应平台的隐私政策。同时，用户又必需使用一些平台，例如微信。平台此时便能够对潜在的风险，通过这些通知性质的隐私政策进行规避。

同时，人工智能公司还会通过“爬虫”程序，从其他的网站、社交平台上，爬取数据，或是获取各种书籍的数据，这一类的数据，很有可能违反了其他网站、社交平台的用户使用政策，或是侵犯了享有书籍著作权的作者的权益。但由于目前监管的宽松，收集、使用这一类数据，已经成为了行业内的“潜规则”。

随着数据收集技术的发展，人工智能公司在保护用户隐私方面面临着越来越多的挑战。用户对于他们的数据如何被收集和使用往往缺乏清晰的了解，这引发了对隐私侵犯的担忧。

5.3 法律和伦理问题

各国法律对数据收集行为提出了规范，但人工智能公司在实践中仍然面临法律界限模糊和伦理边界不清的问题。人工智能公司与个人之间的伦理和法律关系是复杂的。公司需要尊重用户在使用服务时产生的内容，这些内容在一些法律中是受保护的，如欧盟的《通用数据保护条例》(GDPR)[1]，美国的加州消费者隐私法案(CCPA)[2]，中国的《个人信息保护法》[8]等。但公司为了提升用户体验、推出更高质量的服务以提高公司的用户留存率、增加公司的市场份额和竞争力，不可避免地会使用到用户的数据，此时便会出现界限问题。例如，当公司为了改善服务质量而收集用户数据时，他们可能会超出用户预期的范围，从而侵犯了用户的隐私权[9]，使用户产生担忧甚至恐惧。其次，监管部门仍缺乏一个有效手段，对人工智能模型训练过程实施监管，不能很好地对问题的发生进行预防。

另外，在人工智能公司使用到了用户的隐私数据时，是否会隐性地引发基于这些数据的歧视，我们也无从得知。例如，如果雇主或保险公司从医疗数据中了解到敏感的患者信息，若个体被系统认为容易生病或患有治疗成本高的疾病，他们可能不希望雇用该人或为该人提供保险。

6.案例研究

案例研究是一种深入研究特定个体、群体或事件的研究方法，它能够提供对复杂现象的详细见解。在人工智能领域，通过对特定公司或技术的案例研究，我们可以更好地理解人工智能技术的实际应用和潜在影响。本章将通过分析两个案例来探讨人工智能公司的信息收集与使用行为给用户带来的便利和隐私担忧。

6.1 案例一：Google DeepMind 与英国国家健康服务(NHS)的数据共享争议

DeepMind 与 NHS 之间的数据共享争议始于 2015 年，当时 DeepMind 与伦敦的皇家自由医院(Royal Free NHS Foundation Trust)签订了一项协议，用于开发一款名为 Streams 的应用。该应用的目的是帮助医生监测患者的肾脏疾病。

争议的焦点在于，这项合作涉及到了大约 160 万患者的医疗数据，而这些数据被分享给 DeepMind 时，并未充分告知患者，也没有获得他们的明确同意。这违反了患者的隐私权，也违反了当时的数据保护法律。患者的数据包括 HIV 状态、过敏药物记录和历史医疗记录等敏感信息，这些都是进行医疗监测所不必要的数据。

2017 年，英国信息专员办公室(ICO)对这一事件进行了调查，并最终作出裁定，认为皇家自由医院在与 DeepMind 的数据共享过程中未能遵守数据保护法。ICO 指出医院未能向患者提供足够的信息，使他们明白他们的数据如何被使用，也未能就数据共享进行适当的影响评估。

分析：

在此次事件中，涉及到三个主体：医院，公司和患者。医院是患者医疗数据的控制者，而 DeepMind 是患者医疗数据的处理者。医院由于数据处理技术能力上的欠缺，雇佣了 DeepMind 对患者医疗数据进行处理，构建一个医疗辅助系统。然而由于缺乏对于 DeepMind 使用数据的监管和规范，和患者对于自身数据被使用的知情同意，引发了患者医疗数据泄漏的风险，侵害了患者的隐私权。

首先，医院是对患者进行直接护理的主体，无需得到患者明确同意便可利用医疗数据辅助治疗，医院正是在基于这种权利的基础上对医疗数据进行控制和处理。而

DeepMind 作为一家从属谷歌的数据挖掘公司，与患者没有直接护理关系，却将医疗数据存储到自己的服务器中，难以说明 DeepMind 是否拥有对数据进行存储和处理的权利，也难以界定 DeepMind 是否会将医疗数据作为他用。其次，医院和 DeepMind 达成的交易，是在私底下完成的。未经公开招标或发布相关公告。作为数据的控制者，医院并未通过有法律性质的合同或文书，限制和规范数据处理者 DeepMind 的行为和权力，侵害了患者的权利。

其次，从 DeepMind 的角度来讲，它在与医院的交易中，为医院提供了技术上的支持。但是在这个过程之中，缺乏一个有力的监管机构，对 DeepMind 使用患者医疗信息的过程进行监管。而在事件发生之后，DeepMind 对于在过去或实际上如何处理这些数据仍不公开。正如上文所说，DeepMind 使用这些数据的过程是否合法合规，是否只用在了这个医疗系统的建设上，患者和公众无从得知。

对于患者而言，DeepMind 的系统，确实帮助到了小部分的患者。但研究显示有六分之五的患者没有受益于系统，也没有被告知自身的医疗数据被使用，却又承担着个人医疗信息泄漏的风险。[10]

在这个案例中，我们能够得到几点启示：1.在对数据进行使用时，需要告知并征得产生数据的个体的同意，维护个体的合法权益。2.需要有特定机构的监督或具有法律约束力的文件规范数据的使用行为。严格区分数据的控制者、处理者，区分不同角色之间使用数据的权力范围。

6.2 案例二：谷歌 Gemini 模型与百度“文心一言”

谷歌 Gemini 模型是由谷歌公司在 2023 年 12 月 6 日推出的原生的多模态大模型，谷歌称它是规模最大，功能最强大的人工智能大模型。百度的“文心一言”是百度推出的一款人工智能写作助手。它基于百度的大规模预训练模型 PaddlePaddle ERNIE，能够进行深度学习和理解，以生成高质量的文章[11]。

2023 年 12 月 8 日，微博大 V@阑夕及《AI 研究局》等自媒体爆出，如果将以中文询问谷歌的 Gemini 模型它的身份，Gemini 会表示自己是“百度”、“文心一言大模型”。网友进行实验之后，发现 Gemini 确实存在这样的问题。[12]

这么来看，谷歌在训练 Gemini 时，很有可能使用了“文心一言”输出的语料进行训练。然而，有谷歌的支持者认为，这只是大语言模型出现了“幻觉”问题。幻觉问题指的是是大型语言模型，在生成文本时，由于超出了训练数据范围的信息，可能出现的一种生成错误信息的问题。但如果谷歌真的在训练中使用到了百度“文心一言”输出的语料进行训练，而又未征得百度的同意，就有可能构成一种侵权行为。

分析：

首先，我们需要探究谷歌使用“文心一言”生成的语料的版权归属问题。《“文心一言”用户协议》[13]的第 4 条第(10)点提到，“未经百度授权，非法转售或对外提供文心一言服务，或以商业化目的使用文心一言服务。”谷歌 Gemini 与百度“文心一言”作为两个公司分别的大语言模型，拥有几乎相同的使用团体。谷歌公司和百度公司，构成有利益竞争关系的两个主体。同时，谷歌的大模型 Gemini Ultra 版本是付费订阅制的，即是用于商业化目的的。“文心一言”用户协议的第五条第 2 点提到，“百度在本服务中提供的内容(包括但不限于软件、技术、程序、代码、用户界面、网页、文字、图片、图像、音频、视频、图表、版面设计、商标、电子文档等)的知识产权(包括但不限于著作权、商标权、专利权和其他知识产权)属于百度所有，但相关权利人依照法律规定应

享有权利的除外。”，即生成的内容的知识产权归百度所有。如果谷歌真的使用了“文心一言”的输出作为训练数据，那么谷歌就违反了用户协议，同时也侵犯了百度的知识产权。

但这一切都是建立在谷歌在训练中使用了“文心一言”输出的语料进行训练的基础上成立。而事实是否真的如此，是十分难以确定的。由于前文提到的，深度神经网络的黑盒性质，如果不能获取到谷歌训练 Gemini 的训练日志，目前还没有一个有效的手段能够证明在训练中用到了哪些数据。但训练日志对于谷歌而言，可能又属于商业机密而无法公开。况且，谷歌已经根据问题，紧急修复了 Gemini 模型，可所谓的紧急修复，是浮于表面的，是基于规则对特定输出内容进行人为修改，但如果不经过重新训练，修复后的模型只是隐藏了它使用侵权数据的行为，但并未改变它侵权的事实。相反地，在修复之后，更加难以获取谷歌侵权的证据。

在这个案例中，我们能够得到几点启示：1.在人工智能训练中用到的语料，是有可能会通过特殊的提示词，让模型生成出能够辨认训练数据的信息的。这意味着如果模型训练使用了侵犯隐私的信息，那么是存在可能让模型输出同样侵犯隐私的信息的；2.人工智能模型的侵权行为难以确定，公司更是能够通过一些过滤的方法，隐藏自己侵权的行为。

7.数据收集、使用与隐私、知识产权保护的对策

7.1 数据收集与隐私保护的挑战

在人工智能时代，数据是新的石油。然而，随着数据收集、使用的普遍化，隐私、知识产权保护的挑战也日益增加。本章将基于上文提到的人工智能收集、使用数据过程中的隐私、知识产权保护挑战，提出可能的对策。

数据收集、使用带来了多方面的隐私挑战，包括数据泄露、滥用和未经授权的数据共享。这些挑战不仅威胁到个人的隐私安全，也可能导致信任危机，损害公司的声誉和商业利益。尽管各国都在努力制定和完善相关法律法规，但现行的法律体系往往难以跟上技术发展的步伐，导致监管上的空白和执行上的困难。而人工智能的发展，确实是目前世界发展的一个关键点之一，它对推进生产力发展、社会进步有着重要作用。那么如何在当前的技术条件之下，找到一个数据驱动的人工智能发展和个人隐私保护之间的平衡点，就成为了值得讨论的问题。

7.2 数据收集与隐私保护问题的对策

首先，监管部门需要制定相应法律，严格界定个体对于数据的控制权。一个极端是，所有的数据的控制权，归属于产生数据的个体。当涉及到数据的使用时，需要就每一条数据，对每一个拥有数据控制权的个体征询同意。这么做，可以使个体的隐私得到最大程度的保护，但同时也会将许多基于数据的工作扼杀在摇篮中，还会带来大公司垄断数据的局面，十分不利于人工智能技术的发展。另一个极端是，数据被视为是完全可转让的，在这种情况下，个体一旦同意使用条例，就对数据不再保留任何的控制权。在这种情况下，能极大程度让使用大数据的工作摆脱束缚、迅速发展，但同时却又容易引发严重的隐私侵犯问题。所以，一个折中的可能解决方案是，将数据存放于政府部门或慈善信托机构中，由政府或代表患者的委员会，对数据的使用进行授权和规范。在这两个极

端之间，找寻一个合适的平衡点。[14]

同时，也需要有对训练完成后的模型、系统实施审查的能力。由于模型训练过程通常涉及公司的商业机密。政府应该充当审查者的角色，制定相关的法律法规，要求公司保留完整的训练代码、训练日志，以及在指定位置存储使用到的数据。政府应成立相关的部门，对代码、日志和数据进行审查，并监管数据，严格限制数据的复制行为。同时，在模型、系统最终投入使用前，要进行风险评估，并根据风险等级，在后续对模型、系统实施不同程度的监管。另外，在技术上，仍需推进有关解耦智能模型、增强模型可解释性的研究。通过改进模型结构，实现对模型的精细化控制，转变到小数据大任务的智能模式，能够从根源上避免模型中的数据滥用问题和歧视问题。[15]

通过本章的分析，我们可以看到要解决人工智能使用数据过程中的隐私保护问题，需要政府制定相关的法律法规，成立相关的部门、委员会，作为第三方来规范人工智能公司对于用户数据的使用，寻找数据驱动的人工智能发展和隐私、知识产权保护之间的平衡点。另外，也要推动人工智能技术可解释性的发展。

8.结论

在本研究中，我们已经探讨了人工智能公司在信息收集和隐私、知识产权保护方面所面临的挑战，并分析了不同的案例研究以及相关的对策。本章将总结研究的主要发现，并提出针对政策制定者、企业和个人的建议。

本研究表明，人工智能技术的快速发展和广泛应用，带来了生产力的提升。人工智能公司对于用户数据的使用，能够有效改善用户体验，同时增强公司的竞争力。同时，随着人工智能公司收集的数据的增加，隐私保护的问题也变得更加突出。我们发现，由于人工智能系统的特殊性质，尽管存在保护个人隐私的法律法规，在实际操作中仍存在漏洞和不足。

研究指出，主要挑战包括法律法规的滞后性、技术对策的不完善。由于法律法规对于新形式的数据的控制权，没有做出严格规定，容易导致针对数据使用的纠纷。另外，目前欠缺成熟的技术对策，对人工智能公司使用数据的行为进行的审查和监管尚有欠缺。

通过案例研究，笔者给出了一种可能方案：引入第三方，对数据进行存储，对数据的使用行为进行监管和规范。同时，推动有关人工智能可解释性的研究，能够从根源上，有效地缓解数据滥用问题。

9.参考文献

[1]General Data Protection Regulation(GDPR)，https://gdpr-info.eu/
[2]California Consumer Privacy Act (CCPA)，https://oag.ca.gov/privacy/ccpa
[3]The Nature of Statistical Learning Theory, Vapnik, 1999
[4]Induction of decision trees, Quinlan, 1986
[5]Imagenet classification with deep convolutional neural networks, Krizhevsky et. al., 2012.
[6]ImageNet: A large-scale hierarchical image database, Deng et. al., 2009

[7]Training Compute-Optimal Large Language Models, Hoffmann et. al. 2022
[8]《中华人民共和国个人信息保护法》，
http://www.npc.gov.cn/npc/c2/c30834/202108/t20210820_313088.html
[9]Privacy concerns in China's smart city campaign: The deficit of China's Cybersecurity Law, Yang et. al. 2018
[10]Google DeepMind and healthcare in an age of algorithms, Powles, 2017
[11]ERNIE 3.0: LARGE-SCALE KNOWLEDGE ENHANCED PRE-TRAINING FOR LANGUAGE UNDERSTANDING AND GENERATION, Sun et.al 2021
[12]Gemini 自称身份来自百度，谷歌或进行了紧急修复，反映出哪些问题？
https://www.zhihu.com/question/635504283/answer/3330453567?s_r=0&utm_campaign=shareopn&utm_content=group1_Answer&utm_medium=social&utm_oi=94490595935359795 2&utm_psn=1720425595354337280&utm_source=wechat_session
[13]《文心一言用户协议》，https://yiyan.baidu.com/infoUser
[14]Privacy in the age of medical big data, Price et.al, 2019
[15]A Survey on Neural Network Interpretability，Zhang et. al., 2021